

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)





**Canadian Intellectual
Property Office**

Strategis

[Help](#) [What's New](#) [Site Map](#) [Feedback](#) [About Us](#) [Français](#)

GO TO [Main Menu](#)

[Licenses, Legislation and Regulations](#)

Search **Strategis**

**CIPO
Home Page**

**Patents Data
Home Page**

**Trade-marks
Database**

**Search U.S.
and European
Patents**

**Search
Options**

- [Basic](#)
- [Number](#)
- [Boolean](#)
- [Advanced](#)

**Guided
Tour**

Help

- [Content](#)
- [Searching](#)
- [FAQ](#)

Disclaimer

Canadian Patents

11/03/1999 - 10:18:38

(12) **Patent Application:**

(11) CA 2108113

(54) DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID BIOSYNTHESIS

(54) SEQUENCE D'ADN CODANT POUR DES ENZYMES DE LA SYNTHÈSE DE L'ACIDE CLAVULANIQUE

View or Download Images

(72) Inventors (Country):

Jensen, Susan E. (Canada)
Aidoo, Kwamena A. (Canada)
Paradkar, Ashish S. (Canada)

(73) Owners (Country):

Governors of the University of Alberta (The) (Canada)

(71) Applicants (Country):

(74) Agent:

Gowling, Strathy & Henderson

(45) Issued on:

(22) Filed on:

Oct. 8, 1993

(43) Laid open on:

Apr. 9, 1995

(51) International Class (IPC):

C12N 9/00; C12P 17/18; C12N 1/21; C12N 15/52; C12N 15/76

Patent Cooperation Treaty (PCT): No

(30) Application priority data: None

Availability of licence:

N/A

Language of filing:

English

ABSTRACT:

24

ABSTRACT

DNA sequences are provided which encode the enzymes required for clavulanic acid synthesis. A process is provided for producing clavulanic acid in a transformant of a non-clavulanate-producing host.

THIS PAGE BLANK (USPTO)

 Industry Canada	 Canadian Intellectual Property Office	 Strategis
Help	What's New	Site Map
Feedback	About Us	Français
GO TO Main Menu Licenses, Legislation and Regulations		<input type="button" value="Search"/> Strategis

[CIPO Home Page](#)

Canadian Patents

[Patents Data Home Page](#)

11/03/1999 - 10:19:34

[Trade-marks Database](#)

Patent Document Number 2108113 :
DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID BIOSYNTHESIS

[Search U.S. and European Patents](#)

SEQUENCE D'ADN CODANT POUR DES ENZYMES DE LA SYNTHÈSE DE L'ACIDE CLAVULANIQUE

[Search Options](#)

- [Basic](#)
- [Number](#)
- [Boolean](#)
- [Advanced](#)

[Guided Tour](#)

[Help](#)

- [Content](#)
- [Searching](#)
- [FAQ](#)

[Disclaimer](#)

CLAIMS:

19

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. An isolated genomic DNA molecule comprising the nucleotide sequence of Figure 2 (Sequence ID No.:1).
2. An isolated DNA molecule having the nucleotide sequence of nucleotides 2033 to 13636 of Figure 2 (Sequence ID No.:20).
3. An isolated DNA molecule having the nucleotide sequence of nucleotides 109 to 1764 of Figure 2 (Sequence ID No.:21).
4. An isolated DNA molecule having the nucleotide sequence of nucleotides 2216 to 3937 of Figure 2 (Sequence ID No.:22).
5. An isolated DNA molecule having the nucleotide sequence of nucleotides 3940 to 5481 of Figure 2 (Sequence ID No.:23).
6. An isolated DNA molecule having the nucleotide sequence of nucleotides 5654 to 6595 of Figure 2 (Sequence ID No.:24).
7. An isolated DNA molecule having the nucleotide sequence of

(Sequence ID No.:25).

8. An isolated DNA molecule having the nucleotide sequence of nucleotides 7895 to 9076 of Figure 2
(Sequence ID No.:26).

9. An isolated DNA molecule having the nucleotide sequence of nucleotides 9241 to 10908 of Figure 2
(Sequence ID No.:27).

10. An isolated DNA molecule having the nucleotide sequence of nucleotides 10998 to 12296 of Figure 2
(Sequence ID No.:28).

11. An isolated DNA molecule having the nucleotide sequence of nucleotides 12622 to 13365 of Figure 2
(Sequence ID No.:29).

12. An isolated DNA molecule having the nucleotide sequence of nucleotides 13769 to 14995 of Figure 2
(Sequence ID No.:30).

13. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 10.

14. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 11.

15. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 12.

16. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 13.

17. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 14.

21

18. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 15.

19. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 16.

20. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 17.

21. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 18.

22. An isolated DNA molecule comprising a nucleotide sequence encoding the amino acid sequence of Figure 19.

23. An isolated protein having the amino acid sequence of Figure 10.
24. An isolated protein having the amino acid sequence of Figure 11.
25. An isolated protein having the amino acid sequence of Figure 12.
26. An isolated protein having the amino acid sequence of Figure 13.
27. An isolated protein having the amino acid sequence of Figure 14.
28. An isolated protein having the amino acid sequence of Figure 15.
- 22
29. An isolated protein having the amino acid sequence of Figure 16.
30. An isolated protein having the amino acid sequence of Figure 17.
31. An isolated protein having the amino acid sequence of Figure 18.
32. An isolated protein having the amino acid sequence of Figure 19.
33. A recombinant vector comprising a DNA molecule in accordance with any of claims 1 to 22.
34. A host transformed with a recombinant vector comprising a DNA molecule in accordance with any of claims 1 to 22.
35. A host transformed with a recombinant vector in accordance with claim 2 wherein the host is a Streptomyces.
36. A host in accordance with claim 35 which is *S. lividans*.
37. A process for producing clavulanic acid in a non-clavulanate-producing host comprising transforming the host with a DNA molecule in accordance with claim 2 and culturing the host under suitable conditions to produce clavulanic acid.
38. A process for producing clavulanic acid in accordance with claim 37 wherein the host is *S. lividans*.
39. A process for enhancing clavulanic acid production in a clavulanate-producing host comprising

23 transforming the host with a DNA molecule comprising a nucleotide sequence encoding one or more of the enzymes of the clavulanate synthetic pathway.

© *Her Majesty the Queen in Right of Canada, 1999*

[Help](#) [What's New](#) [Sitemap](#) [Feedback](#) [About Us](#) [Français](#) [Top of Page](#)

Canada
<http://strategis.ic.gc.ca>



2108113

1

DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
BIOSYNTHESIS

This invention relates to methods for the production
 5 of the antibiotic, clavulanic acid.

Background of the invention

Clavulanic acid is a broad spectrum beta-lactamase
 inhibitor and is an important antibiotic for the
 10 treatment of infectious diseases. It is produced
 commercially by the gram-positive mycelial prokaryote
Streptomyces clavuligerus, which also produces the β -
 lactam antibiotics penicillin N, desacetoxy
 cephalosporin C and cephamycin C. Until recently,
 15 however, the pathway employed for clavulanic acid
 biosynthesis was much less well understood than the
 pathways leading to these other antibiotics.

Without knowledge of the pathway for clavulanic acid
 biosynthesis, it was not possible to isolate the genes
 20 coding for the key enzymes and to manipulate these genes
 to increase antibiotic yield or permit production of the
 antibiotic in heterologous systems.

One of the earliest enzymes of the pathway to be
 purified and characterised was clavaminic acid synthase.
 25 Two isozymes have now been identified and characterised
 (Marsh et al., (1992), Biochem., vol. 31, pp. 12648-657).

European Patent Application 0349121 describes a DNA
 restriction fragment encoding a portion of the genetic
 information involved in clavulanic acid synthesis but
 30 provides no sequence information.

Until the work of the present inventors, the
 complete complement of genes required for clavulanic acid
 synthesis had not been identified. The present inventors
 have now isolated, cloned and sequenced an 13.6 kb
 35 genomic DNA sequence from S. clavuligerus which codes for
 eight proteins and enables the production of clavulanic



2108113

7

acid by transformants of non-clavulanic-producing organisms.

Summary of the invention

- 5 An isolated genomic DNA molecule is provided comprising the nucleotide sequence set out in Figure 2. A process is provided for producing clavulanic acid in a transformant of a non-clavulanate-producing host.

10 Description of Drawings

The invention, as exemplified by a preferred embodiment, is described with reference to the accompanying drawings in which:

- Figure 1 shows the N terminal amino acid sequence of CLA and the nucleotide sequence of a probe (Sequence ID No.:2) directed to the underlined region of the sequence.

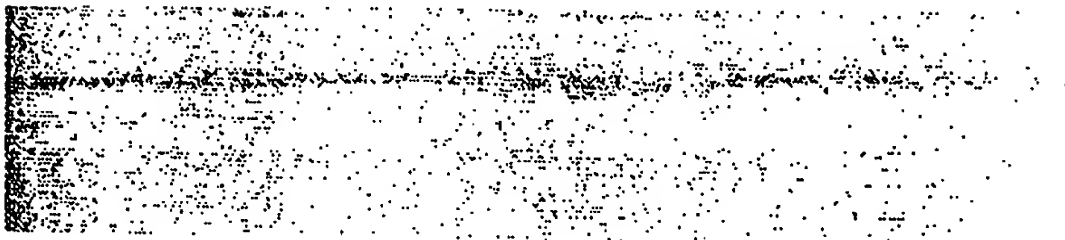
- Figure 2 (2-1 to 2-10) shows the nucleotide sequence (Sequence ID No.:1) of a 15 kb genomic DNA fragment from S. clavuligerus. The sequences of the ten ORFs within the fragment are shown in upper case letters and the intergenic regions are shown in lower case letters. The locations of the beginning and end of each ORF are also indicated directly above the nucleotide sequence. Asterisks above the sequence indicate the EcoRI sites which mark the beginning and end of the portion of the DNA sequence which contains all the genetic information for clavulanic acid synthesis.

Figure 3 shows the location of the open reading frames downstream from pbcG.

- Figure 4 shows a partial restriction map of the DNA sequence of Figure 2 in the region surrounding cla (ORF4).

Figure 5 shows a shuttle vector used for disruption of the ola gene.

- Figure 6 shows a photograph of an agar plate bearing cultures of S. lividans transformants.





2108113

3

Figure 7 shows an alignment of the amino acid sequence of CLA (*S. clavuligerus* CLA) with those of *E. coli* agmatine ureohydrolase (*E. coli* AUH), yeast arginase (yeast ARG), rat arginase (rat ARG) and human arginase (human ARG).

Figure 8 shows a Southern blot of *Nco*I digests of genomic DNA from five presumptive mutants (lanes 1-5) and from wild-type *S. clavuligerus* (lane 6). Panel A : membranes probed with cla-specific probe. Panel B : membranes probed with ter-specific probe.

Figure 9 shows restriction enzyme maps of *S. clavuligerus* DNA inserts in cosmids. A. Restriction enzyme map of cosmid K6L2. B. Partial restriction enzyme map of cosmid K8L2. C. Restriction map of cosmids K6L2 and K8L2 indicating location of pcbc gene in relation to cla. D. The 2.0 kb *Nco*I fragment encompassing the cla gene used in generating nested deletions for sequencing. Abbreviations: Ds, *Dam*HI; B, *Bgl*II; E, *Eco*RI; K, *Kpn*I; N, *Nco*I; S, *Sal*I; and Sm, *Sma*I.

Figure 10 shows the deduced amino acid sequence (Sequence ID No.:3) of ORF1 of Figure 2.

Figure 11 shows the deduced amino acid sequence (Sequence ID No.:4) of ORF2 of Figure 2.

Figure 12 shows the deduced amino acid sequence (Sequence ID No.:5) of ORF3 of Figure 2.

Figure 13 shows the deduced amino acid sequence (Sequence ID No.:6) of ORF4 of Figure 2.

Figure 14 shows the deduced amino acid sequence (Sequence ID No.:7) of ORF5 of Figure 2.

Figure 15 shows the deduced amino acid sequence (Sequence ID No.:8) of ORF6 of Figure 2.

Figure 16 shows the deduced amino acid sequence (Sequence ID No.:9) of ORF7 of Figure 2.

Figure 17 shows the deduced amino acid sequence (Sequence ID No.:10) of ORF8 of Figure 2.

Figure 18 shows the deduced amino acid sequence (Sequence ID No.:11) of ORF9 of Figure 2.



2108113

4

Figure 19 shows the deduced amino acid sequence (Sequence ID No.:12) of ORF10 of Figure 2.

Detailed description of the invention

5 Production of penicillin and cephamycin antibiotics in *S. clavuligerus* starts with the conversion of lysine to α -aminoadipic acid (Madduri et al., (1989), J. Bacteriol., v. 171, pp. 299-302; (1991), J. Bacteriol., v. 173, pp. 985-988). α -Aminoadipic acid then condenses
10 with cysteine and valine to give δ -(α -aminoadipyl)-L-cysteinyl-D-valine (ACV) by the action of aminoadipyl-cysteinyl-valine synthetase (ACVS). ACV is converted by isopenicillin N synthase (IPNS) to isopenicillin N, and, through a series of reactions, to desacetoxycéphalosporin
15 C and ultimately to cephamycin C (Jensen et al., (1984), Appl. Microbiol. Biotechnol., v. 20, pp 155-160).

The ACVS of *S. clavuligerus* has been purified and partially characterized by three separate groups, and estimates of its molecular weight vary from 150,000 to
20 500,000 Da (Jensen et al., (1990) J. Bacteriol., v. 172, pp. 7269-7271; Schwacke et al., (1992); Eur. J. Biochem., v. 205, pp. 687-694; Zhang and Demain, (1990), Biotech Lett., v. 12, pp. 649-654). During their purification, Jensen et al. observed a 32,000 Da protein which co-
25 purified with ACVS despite procedures which should remove small molecular weight components. It has now been found that this protein is not related to ACVS but rather to clavulanic acid biosynthesis. It has been designated CIA.

30 In accordance with one embodiment of the invention, the present inventors have identified, cloned and sequenced the gene (*cia*) encoding this protein.

In accordance with a further embodiment of the invention, the inventors have cloned and sequenced a 15
35 kb stretch of genomic DNA from *S. clavuligerus* which includes the *cia* gene. Within this 15 kb sequence, the inventors have identified an 11.6 kb DNA fragment which,



2108113

5

when introduced into the non-clavulanate producer S. lividans as described in Example 4, enabled that species to produce clavulanic acid. This indicates that the 11.6 kb fragment contains all the genetic information required for clavulanate production.

As will be understood by those skilled in the art, the identification of the DNA sequence encoding the enzymes required for clavulanate synthesis will permit genetic manipulations to modify or enhance clavulanate production. For example, clavulanate production by S. clavuligerus may be modified by introduction of extra copies of the gene or genes for rate limiting enzymes or by alteration of the regulatory components controlling expression of the genes for the clavulanate pathway.

Heterologous organisms which do not normally produce clavulanate may also be enabled to produce clavulanate by introduction, for example, of the 11.6 kb DNA sequence of the invention by techniques which are well known in the art, as exemplified herein by the production of S. lividans strains capable of clavulanate synthesis. Such heterologous production of clavulanic acid provides a means of producing clavulanic acid free of other contaminating clavams which are produced by S. clavuligerus.

Suitable vectors and hosts will be known to those skilled in the art; suitable vectors include pIJ702, pJ0829 and pIJ922 and suitable hosts include S. lividans, S. parvulus, S. griseofulvus, S. antibioticus and S. lipmanii.

Additionally, the DNA sequences of the invention enable the production of one or more of the enzymes of the clavulanate pathway by expression of the relevant gene or genes in a heterologous expression system.

The DNA sequences coding for one or more of the pathway enzymes may be introduced into suitable vectors and hosts by conventional techniques known to those skilled in the art. Suitable vectors include pUC118/119



2108113

6

and pET-11 and suitable hosts include many organisms, including *E. coli* strains such as MV1193 and BL21(DE3).

An oligonucleotide probe based on the N-terminal amino acid sequence of CLA was constructed as shown in Figure 1 and was used to isolate the gene coding for the protein from *S. clavuligerus*, as described in Example 1.

The gene was found to be located in the *S. clavuligerus* chromosome about 5.7 kb downstream of *pbcC*, the gene which encodes isopenicillin N synthase. The gene contains a 933 bp open reading frame (ORF), encoding a protein of molecular weight 33,368. The deduced amino acid sequence was compared to database sequences and showed greatest similarity to enzymes associated with arginine metabolism, notably arginine, ureohydrolase and arginases.

When an internal fragment of the *cla* gene was labelled and used to probe restriction endonuclease digests of genomic DNA from a variety of other *Streptomyces* and related species, evidence of homologous sequences was seen only in other clavulanic acid or clavam metabolite producers, including *Streptomyces jumentinus*, *Streptomyces lipmanii* (7) and *Streptomyces antibioticus*. No cross reactivity was seen to the β -lactam producing species *Novcardia lactamdurans*, *Streptomyces griseus* or *Streptomyces cattleya*, nor to any of a variety of other *Streptomyces* species which do not produce β -lactam compounds, including *S. fradiae* ATCC 19609, *S. venezuelae* 13c and *S. griseofulvus* NRRL B-5429.

Disruption of the *cla* gene, as described in Example 3, led to loss of the ability to synthesise clavulanic acid.

A 15 kb DNA sequence extending downstream from *pbcC* was cloned and sequenced as described in Example 5. The nucleotide sequence is shown in Figure 2. When this sequence information was analysed for percent G + C as a function of codon position (Bibb et al., (1984), Gene, v. 30, pp. 157-166), ten complete ORFs were evident, as



2108113

FIGURE 2 - 6

8181 CGAGGTGCGC GAGGCGCTCG CCGGGGCCCT CGGGCTGCGC GAGGCGGAGA TGCTGATGCG 8220
 8221 CTCCACCGGG BTGATCGGCL GGCAGTACCC GATGGAGAGC ATCGGGGAGC ALCYCAAGAC 8260
 8281 GCTGGAGTGG CCGCGCGGGG AGCGCGGCTT CGACCGCGCG GCGCGCGCCA TCATBAEGAC 8340
 8341 EGACACCGCG CCGAAGGAGG TCGGGTLAG CBTGCGCGGG CCGACCGTTC TGGCATGCG 8400
 8401 CAAGGGCGTC GGCATGCTGG AGCCCGACAT GCGGACGCTG CTGACCTTCT TCGTCACGGA 8480
 8481 CGCCCGGCTG GACCGGCGCG AGCAGGACCG CCTCTTCGCG CGGGTCATGG ACCCGACCTT 8520
 8521 CAACGCGGTC AGCATGAGG CCGACACCTC CACGAGCGAG AGCGCGGTCG TGTCGCGCA 8580
 8581 CCGCGTGGCG GCGGAGGTCG ACCCGGGGGA GTTCGAGGAG GCGCTGCACA CGCGCGCGCT 8640
 8641 GCGCGTGGTC AAGGACATCG CGAGCGACCG CGAGGCGCG GCGAAGCTCA TCGAGGTCCA 8700
 8701 GGTACCGCGC CCGCGGAGG AGCGCGAGCG CAAGCGGTC GCGAAGAGCG TCGTCAAGTC 8780
 8781 CCGCTTGGTG AAGACCGCG TCGACGGCTG CGACCGCAAC TGGGCGCGCG TCGCGATGCG 8820
 8821 GATCGCGAAG TGCTCGGAGC ACACCGACAT CGACCGAGAG CGGGTCAGCA TCGCGTTCGG 8880
 8881 CGAGGTGGAG GTCTATCGCG CGAAGCGCGG GCGCGACCAAG GCGGACGAGC CGCTGCGGCG 8940
 8941 CCGCGTGGCG GAGCATCTGC GCGCGGACGA GGTGCTCATC GCGATCEAGC TCGCATGCG 9000
 9001 GAGCGGGCGC TTCACCGTCT AGCGCTGCGA CCGTACCGAG GCGTATGTCG GCGTCAAGTC 9080
 9081 GAGGTACAGC ACCTGATGCG CGGCGGGGCG CGGGGCGCG GCGCGTTCG CCGTCAAGTC 9120
 9121 CGGTGCGTG TGCTATGCG GCGGTTCGCG CGGTGCGTG CGGGGCGCG CGGGGCGCG 9180
 9181 GCGGGGCGCG GCGGGGCGCG CGGTGCGTG CGGTGCGTG CGGGGCGCG CGGGGCGCG 9240
 9241 ATGAGAGCCA CTCGCTCGAC GACTGCTGAG GAGGCGTTCG ACCTGCGGCT ACCTGCGGCT 9300
 9301 GTGCGGCGCA CCGATGCGCG GCGCGGAGCG CTCGCGCTCG TCGCGACGCA CGACTTCGAC 9360
 9361 TCCTCGACCG CCGCGACACG GTACTAGCGC TACACTGCA ACTTCCTCG GCTCATGCGC 9420
 9421 CGSACGCTCG TCACCTTCA CACCGCGCGC GCGAAGCGCG GCGAGCGGCT CGTGCCEGAC 9480
 9481 CTCGCGGAGT CCGTGGCGCA GTCTTCGAG GAGCGCGCG TCTCGACCTA CGGGTGGCG 9540
 9541 GAGGGCGTGC GTACGAGGA CCGCACGCGC GTCTTCGAG CGGACATCAA GCGCGCATC 9600
 9601 GCGCGGAGCA ACTACGCGAG CGATGCTCG GCGCGGCGTC CGACTTACTT CCGCGACCTC 9680
 9681 CTCGCGACCG AGTACGCGCG CCGCTGCGCG GAGCGCGAGC CCGACGAGCG GTGATGCTG 9720
 9721 GAGACCGCGG AGTACGCGCG CCGCTGCGCG GAGCGCGAGC CCGACGAGCG GTGATGCTG 9780
 9781 CTGCTGCGCA CCGTGGCGTC CCGCGCGCG GAGCGCGAGC CCGACGAGCG GTGATGCTG 9840

Smith, M. L.



2108113

FIGURE 2 - 7

9841 TACCGGCTGC GCGCGTGGC GACCGGCGC TACCGGATCG TCTCGTACAC CCGCGGCGA 9900
 9901 CTGCGCGTTC TGGAGGCCAA TCGGACATCG GACCGCGAGA CCGACCTCGT GCGCGTCCAG 9960
 9961 CCGCGCTTCC GATCGAGGT GCACCTCGGC AAGGATCGC AGGAGGTGGA CCGCATGCTG 10020
 10021 CTGCGCGGCG AGGCTCATGT GAGCTCGCG GCTTCTGCTG TCGAGCGCGC CCGCCAGGAG 10080
 10081 CCGATCTCTG CCGAGCGGGA GCTGCGCGG CACCGCGAGA ACCCGCTGAC CCGCTTCAC 10140
 10141 TCGATCTACT GCTGTCGAG CCGGATCGCC CCGTTCGAGA ATGTCGACTG CCGCGCGGCG 10200
 10201 GTGAGTTCG CACCGGAGA AGCGGCGATG CAGGAGCGT AGCGCGCGCG GCTCGCGGCG 10260
 10261 GACATCGCGA CACCGTCTG GCGCGGAGC CTGCGCGCT ACAAGCGCTT CCGCGGCTAC 10320
 10321 CCGGTCGCGC CCGAGCGGAC CCGCGGCGT GAGCGCGCG CCGCGGAGCT GAGCGTGGCG 10380
 10381 GCGATCGCGC AGCGCTTCG CACCGGATG CCGCGCGCGA AGGAGCGCT CAGGAGTAC 10440
 10441 CCGCGCGCGC AGCGCTGCG CCGCGGCTG GCGCGGCGC GATCGGAGC GAGGCTGCTG 10500
 10501 GACTTCTCGT CCGCGGAGTA CTTCGAGCGC TACCGCGCGT GCGCGGAGTA TCGCGCGAG 10560
 10561 CACCGGATCG GATCATCAT GTTCGCGTGA GCGCGGAGT TCGCGGAGCG ATCGCGCTTC 10620
 10621 CTGCGGAGA TACCGGAGC GCGCGGATG AAGGAGCGC GCGCGGAGTA CAGCGGAGC 10680
 10681 CTGCGGAGC CCGGATCAA CCGCGCTGCG GAGGAGCGC CCGGATCGC CCGCGGCGC 10740
 10741 CCGCGCGCGC GATCTGCGA CCGGATCGC CAGCTCGGA TCGGAGCGC GATCATCGT 10800
 10801 CCGTATCTGT ACCCGCGGTC CCGCTCTAC CCGGAGCGC ACACCGGAA CCGCTTCTG 10860
 10861 ACCGCTCTCT TCGGATGTA CCGGATCGC GCGCTCGCG CCGGATGAGC CCGGATCGC 10920
 10921 GCGCGGCGC GATGAGCGC GCGGAGCGC CCGGAGCGC CCGGAGCGC CCGGAGCGC 10980
 10981 CCGGAGCGC GCGGAGCGC CCGGAGCGC CCGGAGCGC CCGGAGCGC CCGGAGCGC 11040
 11041 CCGGAGCGC CCGGAGCGC CCGGAGCGC CCGGAGCGC CCGGAGCGC CCGGAGCGC 11100
 11101 AACTCGCGCT CTACAGCGC GAGCTGCGC AGGAGCTGCG GCGTGGCGC GCGTGGCGC 11160
 11161 GTGCGCGCGC GCGTGGCGC GAGGAGCGC GCGGCGGAGC ACTGCTCGC CCGGAGCGC 11220
 11221 CCGGCGGTA CCGGCGGTA GCGGCGGTA GCGGCGGTA GCGGCGGTA GCGGCGGTA 11280
 11281 TCGCGGCGC CCGGCGGTA GCGGCGGTA TCGGCGGTA CCGGCGGTA CCGGCGGTA 11340
 11341 TCGCGGCGC CCGGCGGTA GCGGCGGTA GCGGCGGTA CCGGCGGTA CCGGCGGTA 11400
 11401 CCGTCTCGCT ACCGCTCGC CCGGAGCGC TCGGCGGTA GCGGCGGTA GCGGCGGTA 11460
 11461 AAGCGCGGTA GCGGAGCGC GCGGAGCGC GCGGAGCGC GCGGAGCGC GCGGAGCGC 11520

Sin: 4. Lamy



2108113

FIGURE 2 - 0

```

11521 ACGATGCTGG GGTGGACGAG CAGACTCGTG GTCCGGCCCG GCGCTCCCG CAGGGAGCGG 11580
11581 CGCACCAGCG AGGGGTCTTG CACCGCCGGG TGCGTGGGCG AGCCGAGACC TATCGCGTCC 11640
11641 CCGCGGGCGA GGAAGCTCCG GGCACCGAT GCGCCCGTGA TGTGGAGCGG GGTGGGGCGG 11700
11701 GTGAGCCCGG CCAAGTCGAA GACACCTGTC ACCAGGATCT CCGAGCCGGG TCCCCTCTCG 11760
11761 GACACCCAGG TCTCTCCCG CAGATCGGG AGCGAGACCT CCGCCCGGGG GCGCAGCGGA 11820
11821 TGGTCCCGGG GCGAGATCAC CCACAGCGGG TCGTCTAGCA CTCACAGST GCGCAGCGAC 11880
11881 CCGTCCAGGG TGTGCGGGG GACATGGAG CTCCAGGTGT AGCCCGCGTC CACCTCGTAG 11940
11941 CCGCGCACTT GCGCGCGGAC CTGCTCCGGG GCGTCTGCGG GACCGACAG CAGCAGGTCC 12000
12001 AGCGAGCGCG CCGCTCTCTC CACACCTCG TCGAGCAGGG GTTCCCTGGA GACCAAGCGG 12060
12061 AGCAGCTCCG GCGCTCTCAC GCGCTTGGAG CCATGCGCGA AGATATGCTT CCGCGCGGCG 12120
12121 AGGTGCACTT GTGGAAGAA CCGCCCGCGG GCGACCGAGG TCGGGGAGCG CCGCGTGGTC 12180
12181 AGCGGGCGCG TGTGCGGGCT GCGCAGGGTC AGCGGGAGCG CAGCATCCG GTCCAGCGCG 12240
12241 TCGAGTCTBC GCTCCACGGT GCGGTGCGGG ACACCGGTCC GCGGGGCGAC TTCCATGAGG 12300
12301 tctcgcgagt gtccacgag gtccagtaga gacagataga atcggtgag accagcagac 12360
12361 gtcggttctg acccgagaga caatgtcagt tccattttcc gtcagagac gtacagctga 12420
12421 attgtcagaa gtggtctctg caatgtctcg gacagataga taggcagagc cgtctctagg 12480
12481 attctctctg cggggagag gacagagaca ggcagagatg cgggcagct cggggagagc 12540
12541 tcccgagagc gggggagag ggcagagaca ggcagagatg cgggcagct cggggagagc 12600
12601 gacatgtctg gcgagagag gtcagacctg gtcagagatg cgtatgaga tctcgtcgac 12660
12661 ggtcgcgctg tgcgagagcg tcacggcgta gcgagagctc tccgagatgt cctcgagctg 12720
12721 gacgcttcgag atctgagctg tccgctgctc gtacatctcc ttggtgagcg tctgggtgat 12780
12781 gtcgagagcg agctcgtgt cgtgagctcc cggctcgatg acgagagccc gaccccgcg 12840
12841 ctccgtgacc tctgagagca gctgtcgtc gacccgctc acacccagct tctgagctc 12900
12901 gtacagagcg gcttgcgga cggtcagcgg gcccggagtc gaggagatct gaacagagct 12960
12961 gccccttcgag cgcagagat ggggagagcg cggccgagtc atgtacatca gcccagagag 13020
13021 attggtgctg atcatcgag tccagtcggt ggtgtcgagc tctccagcg gcccagagag 13080
13081 catgagctcg gctgtctga cagagatctc gacgagagcg agcgctcga cgtgagagcg 13140
13141 gacgagagcg tccagagct gcccgtcgcg gacgtcgagt tctgagagat gacgcttcg 13200

```

Sim; M. Barany

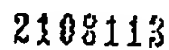


2108113

FIGURE 2 - 9

13201 CCCGCGGCG GICAGCTCGT CACCCAGGGC GCGCAGCTTC TCGACCTGCG GCGCGGCGAT 13200
 13281 GCGCAGGCG GCGGCTTCGG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13320
 13321 GCGGCGGCG ATGAGCGCGA CTTCGCGT GAGTGGGAT GCGATGATG CCGGCGGCG 13380
 13381 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13440
 13441 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13500
 13501 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13560
 13561 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13620
 13621 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13680
 13681 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13740
 13741 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13800
 13801 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13860
 13861 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13920
 13921 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 13980
 13981 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14040
 14041 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14100
 14101 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14160
 14161 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14220
 14221 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14280
 14281 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14340
 14341 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14400
 14401 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14460
 14461 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14520
 14521 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14580
 14581 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14640
 14641 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14700
 14701 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14760
 14761 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14820
 14821 GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG 14880

Sam, 11.11.1999



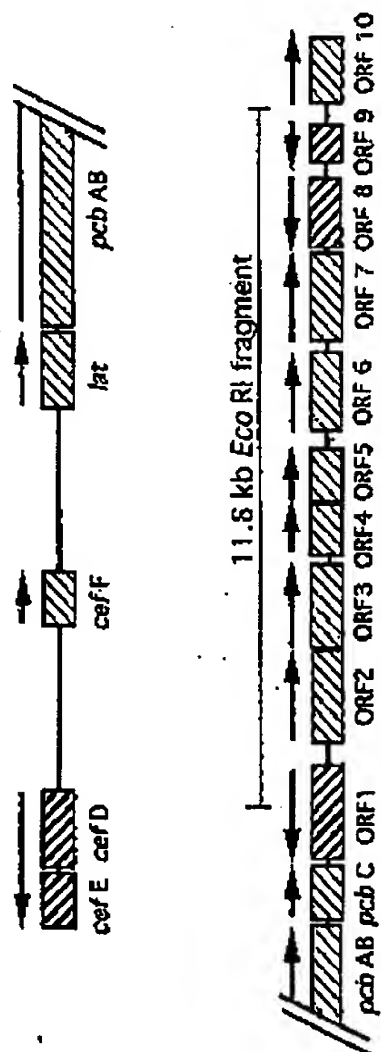
1498	CGGCGCGGTG	CTGCGG	GGA	TGCGCGGTG	CGGCGTGGG	GTGCGCTTG	AGGAGCTGA	14940
1499	CTTCGTCAT	GAGGTGTCA	GTTACGGCT	CGGCGCGCTC	CGGCGACT	GGTGA	gcggc	15000
1500	gtagagcgc	tgacgctgt	ccf.cgcgcg	lcggcctgc	gcgcgctggg	gcgcgtgcgcg		15060
1501	gcacgcgcgc	gcgcgcctc						15079

10 20 30 40 50 60

Sir: J. Barry



2108113



ORF 4 = c/a

FIGURE 3

Simon, M. Luning



2108113

13

program described above. The AVH sequence had previously been aligned with the three ARC sequences (Szumanski & Hoyle (1990), J. Bacteriol., v. 172, pp. 538-547). Identical matches in two or more sequences are indicated with upper case letters.

Example 2DNA hybridization

Genomic DNA preparations from various Streptomyces species were isolated as described by Hopwood et al. (1985). For interspecies DNA hybridization analysis, 2.0 µg amounts of genomic DNA preparations were digested with NcoI for 16h, and electrophoresed in 1.0% agarose gels. The separated DNA fragments were then transferred onto nylon membranes (Hybond-N, Amersham) and hybridized with a cla specific probe prepared by labeling an internal 459 bp BalI fragment (Fig. 1) with (α -³²P)dATP by nick translation. Hybridization was done as described by Sambrook et al., (1989). Hybridization membranes were washed twice for 30 min in 2X SSC; 0.1% SDS and once for 30 min in 0.1X SSC; 0.1% SDS at 65°C.

Sequences homologous to cla in other Streptomyces

Three of six producers of β -lactam antibiotics, S. clavuligerus, S. lipmanii and S. jumonjiniensis showed positive hybridization signals whereas S. cattleya, S. griseus, and M. lactamdurans did not (data not shown). None of the nonproducing strains examined, S. venezuelae, S. lividans, S. fradiae, S. antibioticus and S. griseofuscus gave any signal. All of the streptomycetes that gave positive signals were producers of clavam-type metabolites (Elson et al., 1987).

Example 3Disruption of the genomic cla gene

A 2.0 kb NcoI fragment that contained the entire cla gene was digested at its unique KpnI site and the ends

THIS PAGE BLANK (USPTO)



2108113

14

- made blunt by treatment with the Klenow fragment of *E. coli* DNA polymerase I. A thiostrepton resistance gene (*tsr*), isolated as a 1085 bp *Bcl*I fragment from pIJ702 and cloned into the *Bam*HI site of pUC118 was excised as a
- 5 *Sma*I/*Xba*I fragment and the ends made blunt as above and ligated into the *Kpn*I site of *cl*a. The ligation mixture was introduced into *E. coli* MV1193 and the transformants screened for the presence of the *tsr* gene by colony hybridization (Sambrook et al., 1989).
- 10 Replacement of the chromosomal *cl*a gene by a copy disrupted by the insertion of *tsr*, at an internal *Kpn*I site, was achieved by double recombination. Successful gene replacement was apparent when the 2.0 kb *Nco*I
- 15 fragment which carries *cl*a in the wild type organism was replaced by a 3.0 kb *Nco*I fragment due to the insertion of the 1.0 kb *tsr* gene in the mutants. Four of the five mutants tested showed the expected increase in the size of the *Nco*I fragments, and the larger *Nco*I fragments also hybridized with a *tsr* specific probe. The fifth mutant
- 20 was apparently a spontaneous thiostrepton resistant mutant.

Antibiotic Assay

- The agar diffusion assay was used for determining
- 25 both penicillin/cephamycin and clavulanic acid production. *S. clavuliger* strains to be assayed were grown in 10 ml. amounts of Trypticase Soy Broth (TSB; Baltimore Biological Laboratories) medium with 1.0% starch for 48h. The cultures were washed twice with
- 30 10.3% sucrose and once with MM (Jensen et al. (1982), J. Antibiot., v. 35, pp. 483-490) and the mycelium resuspended in 10.0 mL of MM. Two millilitres of washed cell suspension was inoculated into 100 mL of MM and incubated at 28°C for 48h. The cultures were harvested
- 35 by centrifugation, and the supernatants were assayed for both penicillin/cephamycin and clavulanic acid using



2108113

15

bioassay procedures described previously (Jensen et al. (1982), *supra*).

All of the resulting colonies with disrupted clg genes grew equally well on minimal medium and complex media and produced as much penicillin and cephamycin as did the wild-type, but produced no clavulanic acid (data not shown). HPLC analysis of cell supernatants confirmed the inability of the disrupted clg mutants to synthesize any clavulanic acid (data not shown).

10

Example 4

Protoplast formation and transformation

E. coli competent cell preparation and transformation were as described by Sambrook et al., (1989). Protoplasts of S. clavuligerus were, prepared, transformed and regenerated as described by Bailly et al. (1984), *Bio/Technology*, v. 2, pp. 808-811, with the following modifications. Dextrin and arginine in the regeneration medium were replaced by starch and sodium glutamate respectively. Protoplasts were heat shocked at 43°C for 5 min prior to the addition of DNA. Standard procedures were used for protoplasting and transformation of S. lividans (Hopwood et al. (1985)).

The 11.6 kb EcoRI fragment from K6L2 (Fig. 3) was cloned into the EcoRI site of pCAT-119. pCAT-119 is derivative of pUC119 which was prepared by insertionally inactivating the ampicillin resistance gene of pUC119 by the insertion of a chloramphenicol acetyltransferase gene (Jensen et al. (1989), *Genetics & Molec. Biol. of Ind. Microorg.*, pp. 239-245 Ed. Herschberger, Amer. Soc. Microbiol.). The pCAT-119 plasmid carrying the 11.6 kb fragment was then digested with PstI and ligated to the Streptomyces plasmid pIJ702, which had also been digested with PstI. The resulting bifunctional plasmid carrying the 11.6kb insert was capable of replicating in either E. coli (with selection for chloramphenicol resistance) or in S. lividans (with selection for thiostrepton



2108113

16

resistance). The ligation mixture was transformed to *E. coli*. Plasmid DNA was isolated from several of the chloramphenicol resistant transformants and analyzed by agarose gel electrophoresis to ensure that the proper

5 plasmid construct was obtained. This isolated plasmid material from *E. coli* was then transformed into *S. lividans* as described by Hopwood and transformants were selected by plating onto R2YB medium containing thiostrepton at a concentration of 50 µg/ml.

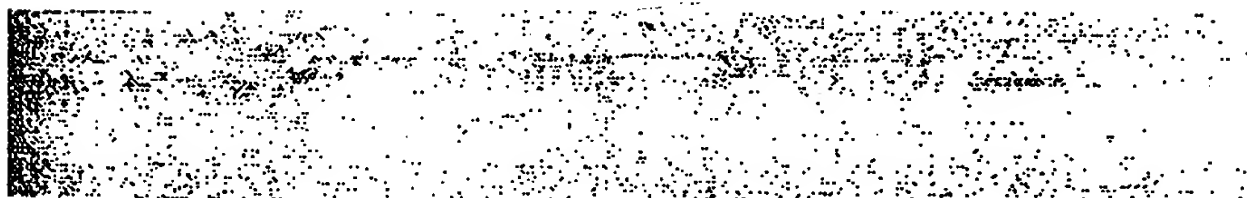
10 Thiostrepton resistant *S. lividans* transformants carrying the bifunctional plasmid with the 11.6 kb insert were patched onto MYM agar plates and allowed to incubate for 48h at 28°C before they were overlaid with molten soft nutrient agar containing penicillin G at a

15 concentration of 1 µg/ml and inoculated with *Staphylococcus aureus* N-2 as indicator organism (Jensen, 1982). (*S. aureus* N-2 was obtained from the Department of Microbiology Culture Collection, University of Alberta. Any organism which produces a β -lactamase sensitive to

20 clavulanic acid may be used as indicator organism.) Zones of inhibition which appeared around the *S. lividans* colonies upon incubation overnight at 30°C were evidence of clavulanic acid production. Clavulanic acid-producing colonies were found amongst these initial *S.*

25 *lividans* transformants at a frequency of about 12%. When plasmid DNA was isolated from one of these clavulanic acid-producing transformants and re-introduced into *S. lividans*, the frequency of clavulanic acid production in these 2nd round transformants was about 40-45%. Figure 6

30 shows a photograph of an agar plate bearing 2nd. round transformants. Zones of inhibition are seen as clear areas in the agar; these appear on the photograph as dark circular areas.





17

Sequencing of 15 kb DNA Fragment

The present invention is not limited to the features
15 of the embodiments described herein, but includes all
variations and modifications within the scope of the
claims.



2108113

18

TABLE 1

ORF #	Start location (bp)	End location (bp)	Length (bp)	Size of ORF (aa residues)
1*	109	1764	1656	552
2	2216	3937	1722	574
3	3940	5481	1542	514
4	5654	6195	542	181
5	6611	7588	978	326
6	7895	9076	1182	394
7	9241	10 908	1668	556
8*	10 998	12 296	1299	433
9*	12 622	13 365	744	248
10	13 769	14 995	1227	409

* Asterisks denote ORFs which are oriented in the opposite direction.



2108113

N-terminal amino acid sequence of CLA	Met Glu Arg Ile Asp Ser His Val Ser Pro Arg															
	Tyr	Ala	Gln	Ile	Pro	Thr	Phe	Met	Arg	Leu	Pro	His	Asp	Asp	(Asp)	
Potential codons (DNA)	TAT	ECT	CAA	ATT	CCT	ACT	TTT	ATG								
	C	C	G	C	C	C	C	C								
		A		A	A	A										
		G			G	G										
Probe made - 24-mer oligonucleotide with 8-fold degeneracy	TAC	GCC	CAG	ATC	CCC	ACC	TTC	ATG								
		G			G	G										
Actual DNA sequence	TAC	GCA	CAG	ATC	CCC	ACC	TTC	ATG								

FIGURE 1

Simon; M. Bunn



2108113

FIGURE 2 - 1

1 10 20 30 40 50 60
 1 gggggggggg cggggggggg gggggggggg cggggggggg cggggggggg cggggggggg 80
 61 gggggggggg gggggggggg cggggggggg cggggggggg cggggggggg cggggggggg 120
 121 gagggttctg aggaacttgc cggggggggg gggggggggg gggggggggg gggggggggg 180
 181 ctccagcagg accgaccagg ccatgtttcc gtcggcctcg tagccgatca tccagggctg 240
 241 cgtcttccgc ggcttctcgc tccgaaactc gggggggggg gtcgttccgc gtcgttccgc 300
 301 cctcagggcc cggggggggg cggggggggg gtcgttccgc gtcgttccgc gtcgttccgc 360
 361 cagcgagtcg acgatggcgc gggggggggg gggggggggg gtcgttccgc gtcgttccgc 420
 421 ggggaccagg accgggctgc taaactcgc cttgttccgc gtcgttccgc gtcgttccgc 480
 481 caccaggggc gacggcctgc ccttggcctg tccggttccg gacggcctgc gtcgttccgc 540
 541 gctgttccgc accgggagcg tccgctcga gtcgttccgc gtcgttccgc gtcgttccgc 600
 601 gatggcgaag gcttccgcgc ccttccgcgc gtcgttccgc gtcgttccgc gtcgttccgc 660
 661 gacgaagcag gcttccgcgc agtgggcga gtcgttccgc gtcgttccgc gtcgttccgc 720
 721 cgtgaactcg tccggttccg cgaagctcg gtcgttccgc gtcgttccgc gtcgttccgc 780
 781 ggcgggcttc tccggttccg tccggttccg gtcgttccgc gtcgttccgc gtcgttccgc 840
 841 ggtgggagcg ggcgggagcg ggcgggagcg gtcgttccgc gtcgttccgc gtcgttccgc 900
 901 ggcgggagcg atgttccgcg tggggggggtc gtcgttccgc gtcgttccgc gtcgttccgc 960
 961 ggcgttccgc ggcgggagcg ggcgggagcg gtcgttccgc gtcgttccgc gtcgttccgc 1020
 1021 cttggcctcg gtcgttccg ggcgggagcg cttggcctcg gtcgttccgc gtcgttccgc 1080
 1081 caggttcagc accgagcggc cgcgggagcg gtcgttccgc gtcgttccgc gtcgttccgc 1140
 1141 ctggagggcc ggcgggagcg tccggttccg gtcgttccgc gtcgttccgc gtcgttccgc 1200
 1201 ctggagggag ttggcgttgc gtcgttccg gtcgttccgc gtcgttccgc gtcgttccgc 1260
 1261 gctctcggcc ggaacgatct ggcgggagcg cttgttccgc gtcgttccgc gtcgttccgc 1320
 1321 cttcgggccc cggagagcgt tgggggagcg gtcgttccgc gtcgttccgc gtcgttccgc 1380
 1381 ggtcatttct accgtgaacg gtatctccac ctcggcctcg gtcgttccgc gtcgttccgc 1440
 1441 ggcgggagcg tccgttccg tggggttgc gtcgttccgc gtcgttccgc gtcgttccgc 1500
 1501 ggcgttccgc ggcgggagcg tgggggagcg gtcgttccgc gtcgttccgc gtcgttccgc 1560

Sim; 4. Summary



2108113

FIGURE 2 - 3

3241 CQTGAGGAC TCAGAGACG CGACGCGTC CTTCGGGCTT AAGCAGCGCC AGGACATCGA 3300
 3301 GCGGCTGCGC GCGCGGATCG CGGAGTTCCT GGCGGACCGG BAGACCTACG AGGACGGCAT 3360
 3361 GCGGCTGCGC CAGGTCATCG ACTCCATCAA CACCGTCATG GAGGAGGCGG CCGAGGCGCG 3420
 3421 CGAGGGGACG ATCGTCTCGG ACATCGGCTT CTTCGCTCAG TACCGTCTGC TCTTCGCGCG 3480
 3481 GCGCGACCGG CCGTTCGCTT TCTTCGCTT GCGCGGCTCG TCGAGCTTCG GCTACGGCAT 3540
 3541 CCGCGCGCGC ATCGGCGCGC AGATCGCGCG CCGCGGACCG CCGAGCTTCG TCATCGCGCG 3600
 3601 TGACGGGCGC TCGCAGCGC ACAGCTCGCA CCGGAGAGCG ACGCGCGCGC TCAACCTGCG 3660
 3661 GATCGTGAGC GTGCTGCTCA ACAACGACAC CAAGGCGCTG ATCGAGCTGT ACCAGAGCAT 3720
 3721 CCGTCACCGC CGCAGCGCGC ACCTGCGCGT CAAGTTCGCG GCGGTCGACT TCGTCGCGCT 3780
 3781 GCGCGAGCGC AACGCTGCTG ACCTGCGCGC CCGCAGCGCG CCGGAGGAGC TCGTCGCGCG 3840
 3841 CCGCGCGCGC GCGCGCGCGC TCGCTGCTCG GTTCCTCATC GAGGTCGCGG TCAACTAGCA 3900
 3901 CTTCAGCGCG GCGCGCTTGG GCGCGCTGAG CATCTGATCA TCGGCGCGCG GGTTCCTCGG 3960
 3961 CCGCGCTTGG GGTTCCTGCG CTTCGCGCGC ACCTGCGCGG GCGCGCGCGC CCGCGCGCGC 4020
 4021 TCGCGCGCGC GCGCGAGCGC CACCGACATC GACAGCGCGC AGGCGGAGCG CCGCTGCGCG 4080
 4081 GCGAGCGCTG TCGAGCGCGC CTGCTGCGCG CCGAGCGCGC GCGGCGCGCG CTCTCTGCGC 4140
 4141 GCGCGCGCGC CCGCGCGCGT GCGCGCGCGT GAGATCTACA ACCGCGAGCG ACTCTCTGCG 4200
 4201 GCGCTGCGCG CCGCGCGCGC GCGCGAGCGC GACCGCGAGC TGCTCTGCGC GGTCTGCGAA 4260
 4261 CCGTATGAGC TCGATGCGCT CCGGCTGCGT AACGCGCGCT TCGCGAGCGC GGTTCGCGAGC 4320
 4321 GCGGAGCGCG TCGTCTGCGC CACCGAGCGC GCGGCTTGGC TGCGGCTGTA CACCTGCTGCG 4380
 4381 GCGCGCGCGC AGCTCGCGCG GTTCAGCGAG GCGAGCGCGC TCGCGCGCGC CCGCGAGCGC 4440
 4441 AACGCGCTTC CCGTCTGCGC CCGCGCGCGC GTCGCGCGCT TGACCGGCTG CTACCGAGCG 4500
 4501 CCGCGCGCGC CCGTCTGCGC CATCGAGCTC GCGTCTGCGC CCGCGCTGCG CCGCGCGCGC 4560
 4561 TGAGCGCGCG GCGCTCTGCG CCGGATGCTG CCGGAGCGCG AGGCGGCTCG CCGCGTCTCG 4620
 4621 GCGCGCGCTG AGAAGCGCGT CCGCGAGCGC GTCAAGCGCG CCGAGAGCGC GTTCGCTGCG 4680
 4681 CTCTCGCGCG GAATCGAGCT CCGCGGCTG GCGGCTGCTG CCGAGCGCGC GCGCGCGCGC 4740
 4741 CTGAGAGCGC TCGCATGCGC CACCGAGCGC TCGAAGCGCT TCGCGAGCGC CCGCGCGCGC 4800
 4801 CTGAGAGCTG TCGGAGCGC GCGCGGCGC ATCAAGCTG CCGAGAGCGC GCTCTGCGCG 4860

Sim; 14. Baumf.



2108113

FIGURE 2 - 4

4861 CAGCTCCCGT ACGGGTGTG GCGCTCCGAG TCGGTGGACC CGSACA?CAT CGAGTACCTG 4920
 4921 LTCCCTCTGA CAGCGCTCTA CCGGGGCTC SAGGGGCGG AGCCCCCAT CCTCACTCGG 4980
 4981 TACGGGCGGG ACATCCCTCT CCGGGGCTG CACCGCGAG ACCGGCTGCC CCGCTGGAC 5040
 5041 ACCGTTCTCG AGCAGGAT GCGCACCTTC GACGGCTGA AGGAGATGTC CLCGGTCTG 5100
 5101 TCCAGGCTGG CCGGGCCTG GACGACCTAC CTCGCTCGG ACCGGGAGGT CCTCGATCTG 5160
 5161 CTGCTCTGCT TGGAGGCTGG GTCAAGCGG CCGCATGGCC GGGACAAAGTG GGTGCTGCGC 5220
 5221 CCGCGGATGG CGACGCTCT CCGGGGCTG ACCTCAATC GGGGCAAGCT CCGCTGCCAC 5280
 5281 GAGCGCTCGG GCACTACGTC CTCGTTCTCC CCGCTGCTG TGGACGAGG TGTGCGCGAG 5340
 5341 GACCGCGTCC ACGAGGCGAA GCGGCAAGTG GTGCGGAGT TGTGCTGCTC CACGGTGGG 5400
 5401 GCGCGACGCG ACCCTCTCGA GGTGGACACC GACGATGTC TGGCTCGCT GGGCGACCGG 5460
 5461 ACCCGCGCGG GGGGCGCTA GCGCGGCGG ggggggccc ggggggccc ggggggccc 5520
 5521 ggggggccc ggggggccc ggggggccc ggggggccc ggggggccc ggggggccc 5580
 5581 ttggggccc tgggggccc ggggggccc ggggggccc ggggggccc ggggggccc 5640
 5641 ggggggccc ggggggccc ggggggccc ggggggccc ggggggccc ggggggccc 5700
 5701 CACCTTCATG CCGCTGCTCC ACGATCCCA GCGCGGCGG TATGAGCTG TGGTATCGG 5760
 5761 AGCCCCCTAC GACGGGCGCA CCACTACCG TCGCGGCGG CGCTGCGG CCGAGGCGAT 5820
 5821 CCGGAGTGG TCGGCTCTA TCGAGCTG CCGATCGAC CCGGCGCGG GCACTTCGA 5880
 5881 CCGATCAAC TGTGCTGAG CCGGCGCAT CAATGAGG CCGTTCGACA TGAACATCG 5940
 5941 GATCGACAG GCGGAGGCG ATCTGCTGG CCGCTGAG GCGGAGGCG CCGTTCGAT 6000
 6001 GATCGGCGG GACGCTCGG TCGGCTGCG CCGCTGCGG GCGGCTGCG AGCAGCACGG 6060
 6061 CCGGCTGCG GTGCTGCG TCGGCTGCG CCGGCTGCG AACCGGCGT TCTAGCGGCG 6120
 6121 CCGGCTGCG CAGGCGCGG CCGTTCGCA CCGGCTGAG GAGAGCTGA TCGACCGCG 6180
 6181 GCGGATGCT CAGATCGGA TCGGCGCGA CAACCGAG CCGGCTGCG TCGACTACG 6240
 6241 CCGGCGCGG GCGCTCGCG TCGGCTGCG GCGGCTGCG GCGGCTGCG GCGGCTGCG 6300
 6301 GCGGCGCGG CTGATCGCG AGAGGCTCG CCGGCGCGG GTGCTGCT CCGTTCGAT 6360
 6361 CCGGCTGCT GCGGCGCGT TCGGCGCGG TCGGCTGCG CCGGCTGCG CCGGCTGCG 6420
 6421 CTGCGCGAG GTGCTGCG TCGGCTGCG CCGGCTGCG CCGGCTGCG TCGGCTGCG 6480
 6481 CCGGCTGAG GTGCTGCG TCGGCTGCG CCGGCTGCG CCGGCTGCG TCGGCTGCG 6540

Simon, M. Bunn



2108113

FIGURE 2 - 5

6541 GATCGGTGCG GAACTGCTCT ACCAATACCC *End of ORF 4--* CCGABCCGAC AGAACCAGT TGTSAagggg 6600
 6601 caatcgtgtc ATGCCCTCTC TGATACGTGA *Beginning of ORF 5--* CTGCACCCCG TACCBCGACG ABCTGCTCC 6660
 6661 GCTCCCTCC GAGCTTCCG ASGTGCCGCG CCGGACCTC CATGCTTCC TCGACCAGGC 6720
 6721 GAAGACGCTG GCGGCCGTC TCCCGACGCG GTGGCCGCG GCTCTCGACA CTTCAACGC 6780
 6781 CGTGGGCAAC GAGGACGGTT ATCTGTGCT GCGCGGCTG CCGGTGACG ACAGCGAGCT 6840
 6841 GCGCGAGACG CGGACCTCCA CCGCGCCCG GCTGACGCG AAGCGGCTG TATGAGAGC 6900
 6901 CATGCTCCG CTGGTCCGCG GCGGCTCCG TCTGACAGG GGTACCAAG AGCTGCGCTC 6960
 6961 GCGGACGGTC TACCACGACG TGTAACGTC GCGCGGCGCG CACTACCTGT CTTCCGAGAC 7020
 7021 CTCCGAGACG CTGCTGAGT TCCACACCGA GATGGGTAC CACATCCTCC AGCCGAACCTA 7080
 7081 CGTCATGCTG GCGTGTCCG GCGCGGACCA CAGAACCGG CCGGAGACCG TGGTCCGCTC 7140
 7141 GGTCCGCAAG GCGCTGCCG TCGTGGACGA GAAGACCGG GCGGCTCTCT TCGACCACAA 7200
 7201 GGTGCCCTGC TCGCTGACG TGGCTTCCG CGGCGGCGTC GACGACCGCG GCGGATCCG 7260
 7261 CAACGTCAAG CCGCTCTACG GCGACGCGAA CGACCGTTC CTCGGGTACG ACCGCGAGCT 7320
 7321 GCTGGCGCGG GAGSACCCCG CGGACAGGGA GCGGCTCGCG CATCTGTCC AGGCGCTCGA 7380
 7381 CGATGTGACC GTCGGGGTGA AGCTGCTCC CCGTACGTC CTCATCATCG ACAACTTCCG 7440
 7441 CACGACGCGC GCGCGGACCG CATTCTCCG CCGCTGGAC GCGAAGGACC GTTGGCTCGA 7500
 7501 CCGCGCTTAC ATCCGACCG ACCGCAATGG ACAGCTCTCC GCGGCGGAGC GCGCGGCGCA 7560
 7561 CACCATCTCG *End of ORF 5--* GCGGCTGAGC cgggtcccc ggggacctgg gcccggggg 7620
 7621 cgggacgggg tccgggtctt gcccctctcc cggcgcgcg ggtgaggggg cgggacctt 7680
 7681 tgtgcgggtt ggcgtgcgtc ctgcgggggt gcccggggcg gggggacggc gggggtgcc 7740
 7741 gggggcgggg tgcgtgcgc cgcacgtggg tgcgtgacg cactcgtgt ggcgtgcgc 7800
 7801 ccccggtgca laatttgca actctatggg aaatgagca ggtgagggc ggtgagggc 7860
 7861 tgcgggtgca attacgtga cgggagggc *Beginning of ORF 6--* tgcATGTCC CACAGCACAC CGAAGACGCG 7920
 7921 CCGGCGATTC GTGTGACAC CCGCGCCCGT GCGGCTGCG GACGACGCGC GTCACGACTT 7980
 7981 CACGCTCCTC GCCTCACCG CCGCGGCGAC CGTGAGCGCG CTCCTACCC GTCCCGCTT 8040
 8041 CCGCGGCGCG AGCGTGTGC TGTGCGGGA GCGGCTGCG GACGCGGAGC CCGCGGCTGT 8100
 8101 GGTGCTGCTG GCGCGAACG CGAATCTCC GACCGGCTG GAGGCGGAGG AGAACCGCG 8160

Simon; 14. Lanning



2108113

FIGURE 2 - 5

				End of ORF 4--	
6541	GATCGGTCG	GAATGCTCT	ACCAGTACGC	TCGAGCCGAC	AGAACCAGT
6601	ccatcgctgc	ATCGGCTCTC	TGATAGTGA	CTGCAACCCG	TACCGCGACG
6661	CTCTCGCTTC	GAGLTCGCG	AGGTGCCGCG	CGCGGACCTC	CAAGGCTTCC
6721	GAAGACGCTG	CCCGCCGCTC	TCCCGGACGG	CTTGGCCGCG	GTCTCTGACA
6781	CGTGGGCAAC	GAGGACGGTT	ATCTGCTGCT	GTGCGGCTCG	CCCGTCGACG
6841	GCCTCGAGACG	LCGACCTCCA	CCCCCGCCCC	CTGTGACGCG	AAACCGCTGG
6901	CATGCTCGCG	CTGGCCGCGG	CCCGGCTCGG	TCTGCGACAG	GGGTACGAGG
6961	GCGCACGGTC	TACCGAGACG	TGTACCCGTC	GCCCGCGCGC	CAGTACCTGT
7021	CTCGGAGACG	CTGCTGGAGT	TCCACACCGA	GATGCGGTAC	CACATCTCTC
7081	CGTCATGCTG	GGCTGCTCCG	GGCGGAGCCA	CGAGAACCGG	CGCGAGACGC
7141	GGTCCGCAAG	GGCGTGCCTC	TGCTGACGA	CAAGACCCCG	GGCCGTCTCT
7201	GGTCCCTTGC	TGCTGAGAGG	TGGCTTCCG	CGGCGGGGTC	GACGACCCGG
7261	CAACGTCAAG	CGGCTCTACG	GGGACGCGAA	CGACCTGCTC	CTCGGCTACG
7321	GTGCGGCGCG	GAGGACCTCG	CGGACAGGGA	GGCGCTCGCC	CATCTGTCCG
7381	CGATGTGACG	GTGCGGCTGA	AGCTGCTGCT	CGGTGACGTC	CTCATCATCG
7441	CACGACGACG	GGCGGACGCG	CGTCTCTCCG	CGGCTGCGAC	GGGAAGGACC
7501	CGCGCTCTAC	ATCGCGACCG	ACCGCAATGG	ACAGCTCTCC	GGCGGCGAGC
7561	CACCATCTCG	TTCTGCGCGG	GGCGGTGAGG	CGGGCTCCCG	GGGGGCTGGG
7621	GGGGGCGGGG	TCGGGCTCTC	GGGGGCTGCG	CGGGGCTGGG	GGGGGCTGGG
7681	TGTGCGGGGT	GGCGTGGCTC	CTGCGGGGT	GGCGGGGGGG	GGGGGGGGGG
7741	GGGGGCGGGG	TGCGGTGGCG	GGCGGTGGGG	TGCTGTGACG	CGGCGGTGCT
7801	GGGGGCTGCG	GGCGGTGGCG	GGCGGTGGGG	GGCGGTGGGG	GGCGGTGGGG
7861	TGCGGTGGCG	GGCGGTGGCG	GGCGGTGGGG	GGCGGTGGGG	GGCGGTGGGG
7921	GGGGGCTGCG	GGCGGTGGCG	GGCGGTGGGG	GGCGGTGGGG	GGCGGTGGGG
7981	GACGCTCTCG	GGCTGACCGG	CCCCGGCCAC	CGTGAAGGCG	GTCTGACCGG
8041	CGCGGCGCGG	AGCGTGTGTC	TGTGCGGGGA	GGCGGTGGGG	GACGCGCGAG
8101	GGTGTGCTCG	GGCGGCGAGG	CGAATGCTCG	GACGCGCGAG	GACGCGCGAG

Living; 4. Burning



2108113

FIGURE 2 - 6

8181 CGAGGTGCGG GAGGCGGTGG CCGGGGCGCT CGGGCTGGCG GAGGCGGAGA TGCYGATGCG 8220
 8221 CTCGACCGGG GTGATGCGCG GGCAGTACCG GATGGAGAGC ATCGGGGAGC ACCYCAAGAC 8280
 8281 GCTGGAGTGG CCGCGCGGGG AGGGCGGCTT CGACCGGGGG GCGCGCGCGA TCATGACGAC 8340
 8341 CGACACCGCG CCGAAGGAGG TCGGGTGAAG CGTGGGCGGG CCGACCGTGG TGGGATGCG 8400
 8401 LAAGGGGCTG GGCATGCTGG AGCGCGACAT GCGGACGCTG CTGACCTTCT TCGGACGCGA 8480
 8481 CGCGCGGCTG GACCGGGCGG AGGAGGAGCG CCTCTTCTCG CGGGTATGG ACCGACCGTT 8520
 8521 CAACGGGCTG AGCATGAGA CCGACACCTC CACGAGCGAC AGGGCGGCTG TGTTCGCGAA 8580
 8581 CGCGCTGGCG GCGGAGGCTG AGCGCGGGGA GTTCGAGGAG CGGTGCGACA CGCGCGCGCT 8640
 8641 GCGCGTGGCT AAGGACATCG CGAGCGACGG CGAGGGCGCG GCGAAGCTGA TCGAGGTCCA 8700
 8701 GGTACCGCGG CCGCGGAGCG AGCGCGAGCG CAAGCGGCTG GCGAGAGCTG TCGTCAACTC 8780
 8781 CCGCTTGGTG AAGACCGCGG TCGACGGCTG CCGCGCGAAC TGGGGCGCGG TCGCGCATGG 8820
 8821 GATCGGCAAG TCGTGGGAGG ACACCGAGAT CGACCGAGAG CGGGTGAAGA TCGGCTTGG 8880
 8881 CGAGGTGGAG GTCTATCGCG CGAAGGCGCG GCGCGAGCGA GCGGAGCGCG CGGTGCGGGC 8940
 8941 CGCGCTGGCG GAGCATGCTG GCGCGGAGCG GGTGCTCATC GCGATCGAGC TCGCGATCGC 9000
 9001 CGAGCGGGCG TCGACCGCTT AGGGCTGGGA CCGTACCGAG GCGTATGCTG GCGTGAAGTC 9080
 9081 GAGTACAGCG ACCTGAAGCG GGGGCGGGGA CGGGGCGCG GCGCGTGGCG GTGCGGCTG 9120
 9121 CGGTGGGCTG TGGTATGCG GCGGTGCGCG GCGGTGCGCG CGGGGCGCG GCGGCGCGCG 9180
 9181 GCGGGGCGCG GCGGGGCGCG GCGGTGCGCG GCGGTGCGCG GCGGTGCGCG GCGGTGCGCG 9240
 9241 ATGGAGAGCG CTCGGTGGAG GAGCGTGGAG GAGGGCTTGG AGCGCGGGGT AGCGGGAGTG 9300
 9301 GTGGCGCGCG CGGAGCGCGG GCGGGGCGCG CTGGCGCTGG TCGCGAGCGA CCGCTTGGAG 9360
 9361 TCGCTGAGCG CCGGCAACAG GTACTAGCGG TACAGCTGGA AGTTCCTCG GCTCATEGCG 9420
 9421 CGGAGGCTGG TCAGCTTGGG CACCGCGCGG CGCAAGCGCG GCGAGGCGGT CGTGCGCGAG 9480
 9481 CTGGCGGAGT CCGTGGCGGA GTCTGCGAG GAGGGCGGGG TCTGGAGCTA CCGGCTGGCG 9540
 9541 GAGGGGCTGG GGTACGAGGA CGGACGCGCG GTGCTGCTGG CGGACATCAA GCGCGCATC 9600
 9601 GCGCGCGCGA ACTACGGCGG GGTATGCTCG GCGCGCGGTC CGGCTACTT CCGCGAGCTC 9680
 9681 CTGGCGAGCG AGTACGCGG CCGCTGCGCG GAGCGCGAGC CCGAGGAGCG GTGAGGCTG 9720
 9721 GAGACCGCGG AGGAGCGGAG GCTGCTCTTC CCGCTGCGGG AGCGGCTCGG GGGATGGAT 9780
 9781 CTGCTGCGCG CCGTGGCGTC GAGGCGCGCG GTGCGCGCGG ACCCGGAGAG CGGCGCGGAG 9840

Simon, M. L.



FIGURE 2 - 7

9841	TACCGGCTCG	GGCTCGTGGC	GACCGGCGCG	TACCGGATCG	TCTCG/ACAC	CGCGGGCGAA	9800
9901	CTGGCCGCTC	TGGAGGCCAA	TCCGCACTCG	GACCCCGAGA	CGGACCCGCT	GGCGCTCCAG	9900
9981	CGCGCCCTCC	GGATCGAGST	GCACCTCGGC	AAGSALCCGC	ACGAGGTGGA	CGGCATCTCT	10020
10021	CTGGCGGGCG	AGGCTCATST	GGACCTCGCG	GGCTTCGCTG	TGCAGCTCGC	CGCCGAGGAB	10080
10081	CGCATCTCTC	CGGAGCGGGA	GCTGCGGGBB	CAGCGGGACA	ACCGGCTGAC	CGGCTTCACG	10140
10141	TGGATCTACT	GCTGTCTGAG	CGGATCGGCC	CGCTTCGACA	ATGTGCACTB	CGGCGGGGCG	10200
10201	GTGCACTCTC	CGACCGACAA	AGCGGCGCAT	CAGGAGGCGT	ACGCGGGGCG	CGTGGGCGGG	10260
10261	GACATCGCGA	CGACCTTGCT	GGCGGCGGCG	CTCGACCGCT	ACAAGCACTT	CGACCGCTAC	10320
10321	CGGCTCGGCC	CGGAGGCGAC	CGGCGACCTG	GAGGCGGCTC	GGCGCGGAGCT	GAGGCTGGCG	10380
10381	GGGATCGCGG	ACGGCTTCG	CACCGGGAAT	CGCGGCGCGA	AGGACCGGCT	CAGGAGGATC	10440
10441	CGGCGCGGCG	AGGCGCTGCG	CGCGGGGCTC	GGCGGCGCTG	GCATCGGAGC	GGAGGCTGCT	10500
10501	GACTTCGCTG	CGGCGGACTA	CTTCGAGCGC	TACGCGGCGT	GGCGGCGACTA	TCGCGGCGAG	10560
10561	CACCGGATCG	GGATCATCAT	GTTCGGCTGG	GGCGGCGAAT	TCCCGGAGCG	ATACGGCTTC	10620
10621	CTCGACGAGA	TCACCGAGCG	GGCGCGGATC	AAGGAGCGCG	GCACCGAGAA	CATCGGCGAG	10680
10681	CTGGAGGAGC	CGGAGATCAA	CGCGCTGCTG	GATGAGGCGG	CGGAGTGGCG	CGACCGGCGG	10740
10741	CGGCGCGCGG	AGATCTGGGA	CGGATCGGAC	CAGCTCGGGA	TGGAGCGAGC	GGTCATCGGT	10800
10801	CGGATCTGCT	ACCGCGGCTC	CGGCTCTTAC	CGGCGGCGCG	ACCGCGGCAA	CGGCTTCGCT	10860
10861	ACCGGCTGCT	TGGGATGATA	CGGCTAGCTG	GGGCTCGGCG	CGGAGTGGAG	CGGCTTCGCG	10920
10921	GGGCGGGGCG	GGTATGCGG	GGGCGGGGCG	GGGCGGGGCT	CGGCGGGGCG	CGGCTTCGCG	10980
10981	GGGCGGGGCG	GGGCGGGTCA	GGGCGGATC	GGGCGGGGCG	CGGCGGAGCG	GGGCGGGGATC	11040
11041	GGGCGGGGCG	CGTGGGCGAG	GGGCGGCTCG	ACGCTCGGCG	AGGCGGAGAGC	GGGCTTCGCG	11100
11101	AACTCGGCGT	CGTACGCGCG	GGGCTGGGCG	AGGAGCTGCG	GGGCTGGGCG	GGTLAGGCTG	11160
11161	GTGCGCGCGG	GGTGGGCGAB	CAGGAGGCGG	GGGCGGAGG	AGTGGTTCAG	CGGCTGAATC	11220
11221	CGGCGGGTGA	CGGCGGAGTG	GCTGATGAG	AGGAGCGCGG	CGGCGGGGCT	GATGCTGCGG	11280
11281	TGGCGGGCGA	CGGCTCGGAG	CAGATCGAGA	TCGCTGAGAT	CGGCTTCGCG	GGGCTTCGCG	11340
11341	TGGCGGGGCA	CGGAGGCGTG	GTGCGGCTCG	GGGCGGAGG	GGGCGGCGCT	CGGCGGGGCTG	11400
11401	CGGCTCGGCT	ACGCTGCGCG	CGGAGGCGG	TCGCTGAGCA	GGTGGGCGTG	GTGCTGCGCG	11460
11461	AACTCGGCGG	GCTGAGGCTG	GGGAGGAGC	GGGCGGAGG	GTGCGGCGTG	CGGCGGGGCG	11520

Swim; of. bearing



FIGURE 2 - 0

[illegible]

Swine; 4. Barony



2108113

FIGURE 2 - 9

13201 CCCGCGCGCG G/CAGCTCGT CACCCAGGCG GCGCAGCTTC TCGACCTCGC GCGCGCGAT 13200
 13201 GCGCAGCGCG GCGCGCTCGC GCGCGAGGCG GCGCGCTCGC GCGCTCGCGA TCGCGGAGCT 13200
 13321 GCGCGCGCGT ATGAGCGCGA CTTTCCCTCG GAGTCCGCGT GGCATCATT CCGCGCATG 13380
 13381 GCGTGGCGT CCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13440
 13441 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13500
 13501 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13560
 13561 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13620
 13621 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13680
 13681 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13740
 13741 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13800
 13801 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13860
 13861 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13920
 13921 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 13980
 13981 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14040
 14041 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14100
 14101 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14160
 14161 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14220
 14221 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14280
 14281 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14340
 14341 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14400
 14401 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14460
 14461 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14520
 14521 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14580
 14581 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14640
 14641 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14700
 14701 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14760
 14761 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14820
 14821 GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT GCGTGGCGT 14880

Scanned by



2108113

FIGURE 2 - 10

14881 CGCCBCGGTG CTGCGTGGG TGCCTGGTGC CCGGCTCCG GTGCCCTTG AGGAGCTGGA 14940
 14951 CTTCGGTCAT GAGGTGTCCA GTTACGGGCT CCGCGCCCTC CCGTGACCT GGTGAGCGGC 15000
 15001 gtagagcggc tgacggtgt cctcagcgcg taggcctgct gggcggtggg ggcgtgogcg 15060
 15081 ggcacggccc cagagatct 15079

| 10 | 20 | 30 | 40 | 50 | 60

Simon; H. Baumf



2108113

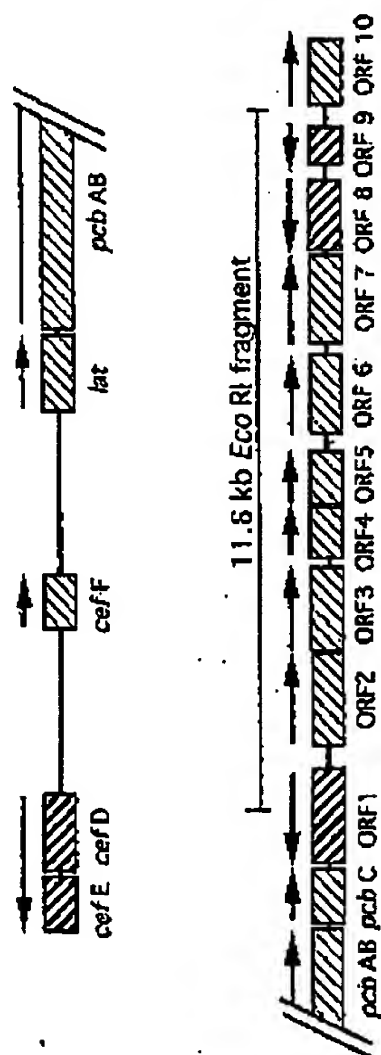
ORF 4 = *cia*

FIGURE 3

Simon, M. G. G.



2108113

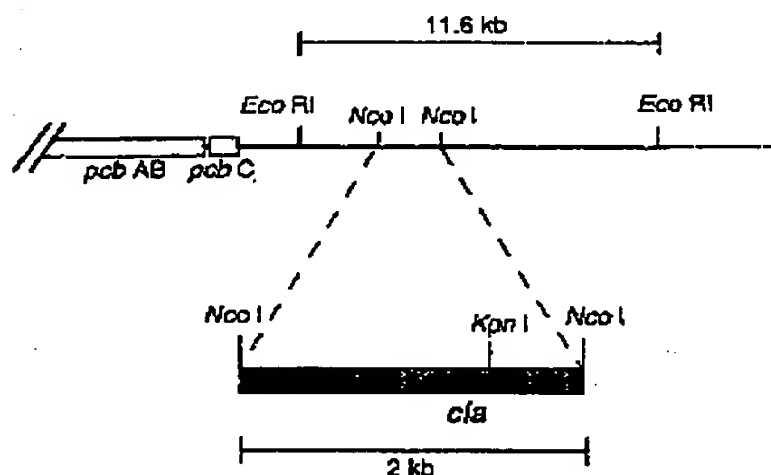


FIGURE 4

Simon, M. Baumf



2108113

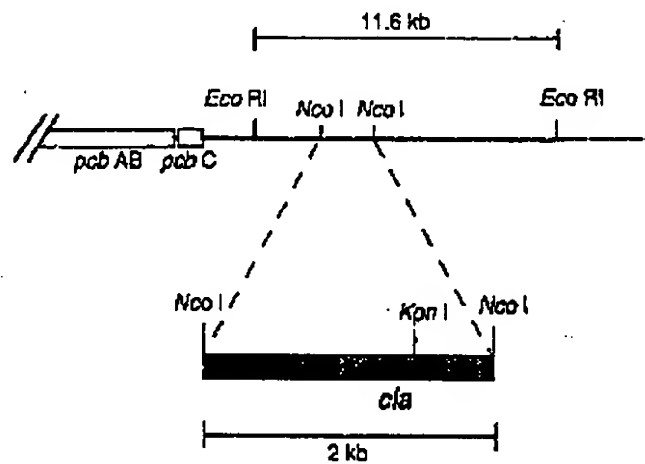


FIGURE 4

Sim; M. Baum



2108113

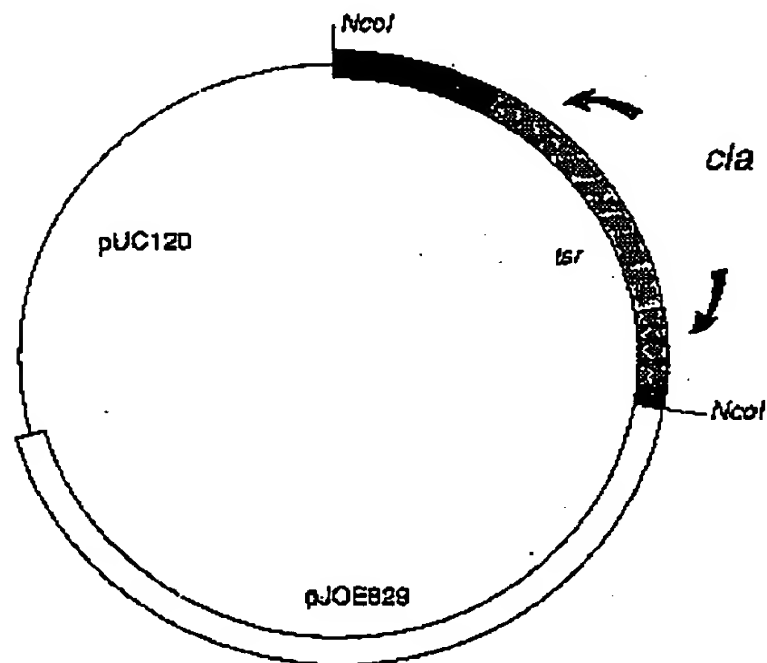


FIGURE 5

Simon, H. Bunn



2108113

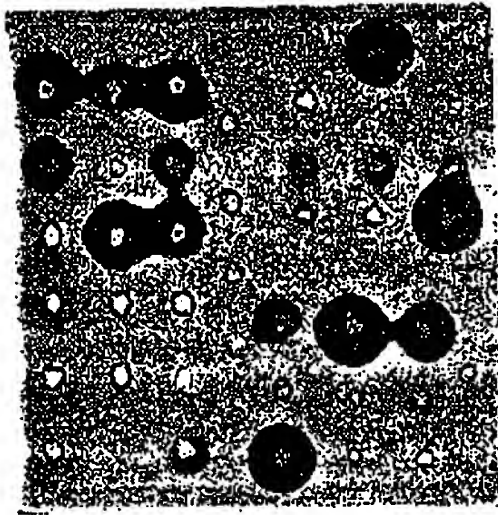
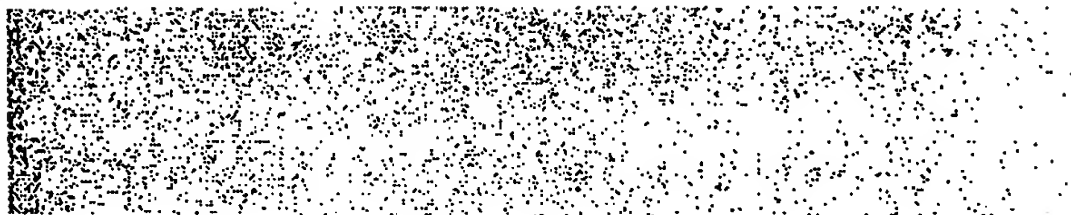


FIGURE 6

Scanned by M. B. B. B.



2108113

```

1
S. Cl. CLA 1 veridshvspryaq1ptFmRLPhdpQPrgyDV--YvIGaPyDggTSyRpGARfGPqAIR 60
E. co. AUH NSTIGhQYdNsivSnaifGfRLPmrfQPydsDaa-vitGvP fDwaTSgRasGRhGPaaIR
yeast ARG MoI-GphY-NyyKtRelsivIAPFsgGgkIGYEKGfymIKhGL-qtalodlgwetoLE
rat ARG MS-----aKpkp1e1IGAPFSKGQPRGCVKGPaaLRKAGL-----vE
human ARG MS-----aKSRtIG1IGAPFSKGQPRGCVKGPvLRKAGL-----LE

61
S. Cl. CLA 81 saSgI(hgvgdRgPgtFDI---fNcyDqGDiN1tpfDmnia1at0aHISgLLKANaaf 120
E. co. AUH qvStnl-awehnRfPmrdmre1NYVDGDIvya1gDarEmSEKLOAKaekLLaAGkcm
yeast ARG psndao-qfYgKikmakdattgssvMidGVKakRad1VGEAtkIvynsYKYYqANRfp
rat ARG KLKEtE-ynV-rDhGDLufvDvPNDSPFO1YKKNPRs--VCKANEQLAAVFAetqKNGtIS
human ARG KLKEqE-cdV-KDyQDLpfad1PNDSPFQ1YKKNPRs--VCKASEDLAqVAAqYKKNGRIS

121
S. Cl. CLA 121 LmIGGDHSL1vuaLRAYAaahGdLAVYHIDAHsOTNpafyGgryhHQTaFrhgideKk10 180
E. co. AUH LatGGDH1v1IpiLRaAKh1fkaALVHfDAHTDTyan--GcofdHGTnFytapkeGLID
yeast ARG LtLGGDH31A1Q1vSA1y1dkyPDaGLIMIDAHaD1KTI--eaTpsGNLHGPVSLmg1n
rat ARG vYLGGDH5m1GS1SaHARYHPDLav1WYDAHTD1NTP--LTTaSGNLHGOPYaFLKEL
human ARG LYLGGDHSLA1GS1SghARYHPDLGY1WYDAHTD1NTP--LTTaSGNLHGOPYaFLKEL

181
S. Cl. CLA 181 PaamVOIGIRGNXPZPD9LdyarhGveVutAdefgalgVggtadL1rakY----- 240
E. co. AUH PnhsVQ1g1rt-----afkdnG1tV1dAcqvnDreYddv1aayka1v-----
yeast ARG Kdvphcpa1k-----WYpgn1SpKk1a1y1GLKDVOaSEkk1LXa1G1aafSNyhyD
rat ARG KCK1PDVPGF5-----WYTPC1SAKD1Y1GLRDVDPGEHY11KTLG1KYFSHTeVO
human ARG KCK1PDVPGF5-----WYTPC1SAKD1Y1GLRDVDPGEHY11KTLG1KYFSHTeVO

241
S. Cl. CLA 241 -----BqRPYYV3u01DvYDPAFAPGTGTpocGL1SKEVLdL1R 300
E. co. AUH -----GdnPYYL1fD1DcLDPAFAPGTGTpV1GGLYSdra1KLYR
yeast ARG KyG1a1v1Emamkavhpaetn0ag1mcSyOVDQVOP1y1PATGT1PVRaGL11REG1FLYE
rat ARG KLQ1QKYME--ET1SYLLGRKKRP1HLSFDVGLDPvFTPATGT1PVYVGLTYREG1YTE
human ARG rLG1QKYME--ET1SYLLGRKKRP1HLSFDVGLDPvFTPATGT1PVYVGLTYREG1YTE

301
S. Cl. CLA 301 cy-g0LkpVG1DYMEVap1YDhggit1a1-----1a1a1gaELLYayArohrTq1a 360
E. co. AUH gL-KDLN1VgmdVVEYap1YDhggit1a1-----AAAt1ALEmLY1aAak1ga
yeast ARG rLaesGNL1aL0VVEcNPdLa1hd1hVant1aoccat1ArcALGet11
rat ARG E1YKTGLLSGLD1MEVNP1LGKTPEEYTRTYNTAVa1TLacFGLaREGNHKP1TOY1KPPK
human ARG E1YKTGLLSGLD1MEVNP1LGKTPEEYTRTYNTAVa1TLacFGLaREGNHKP1TOY1KPPK

```

FIGURE 7

Sim; 4. Barany



2108113

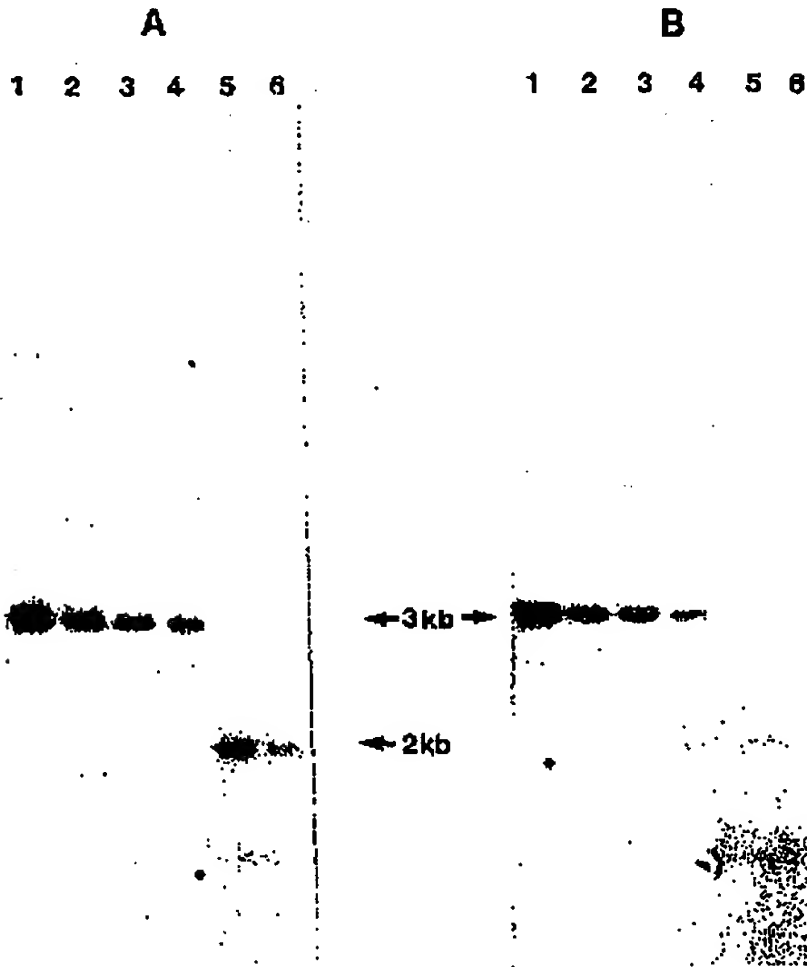


FIGURE 0

Sim; M. Baum



2108113

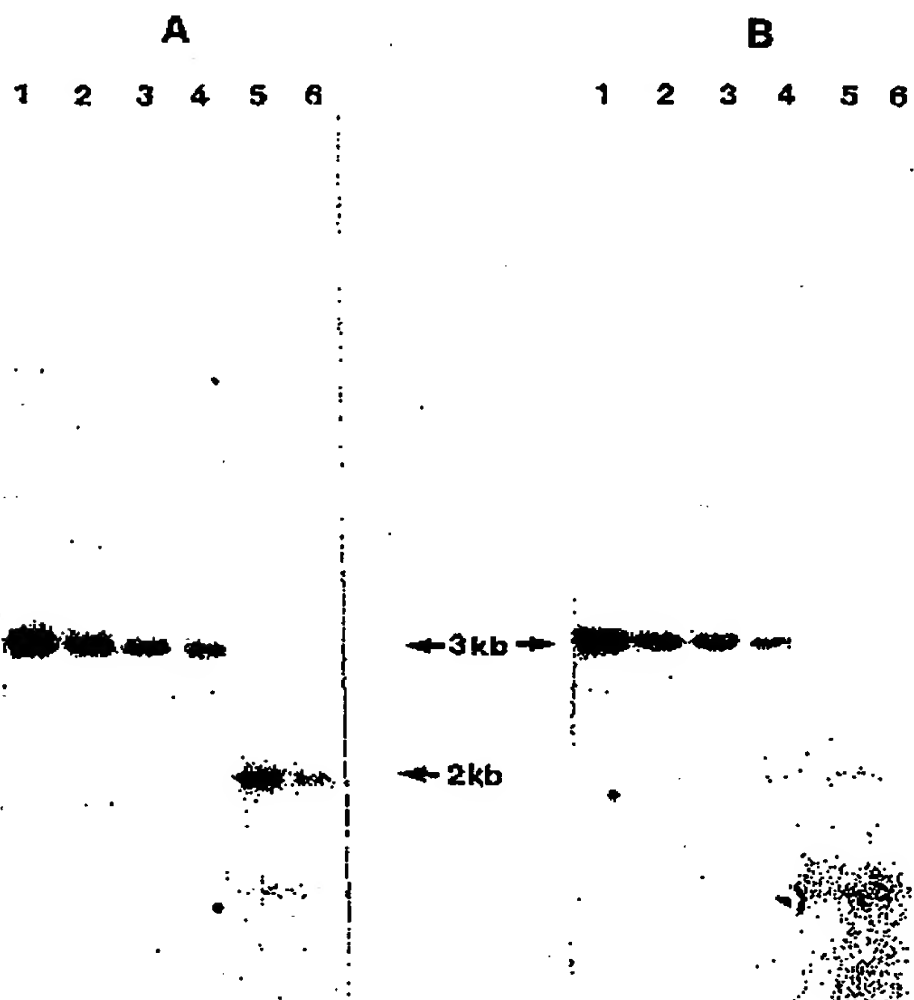


FIGURE 0

Simon, M. Banning



2108113

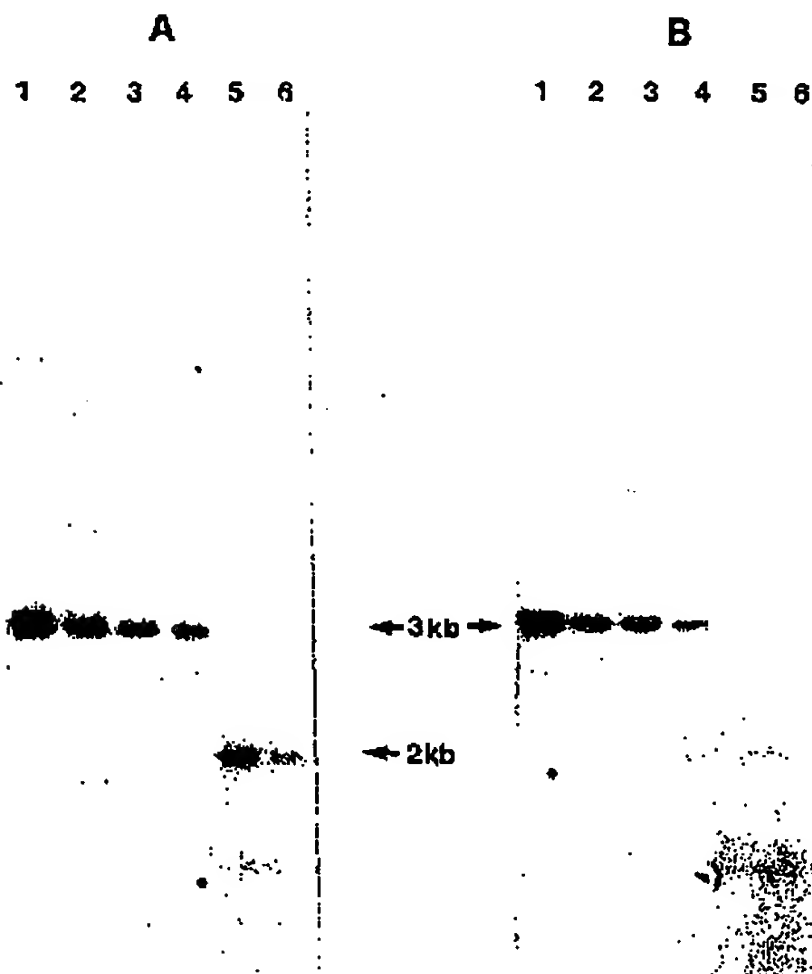
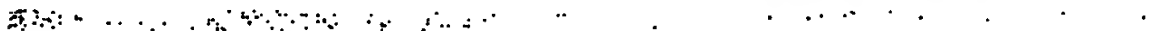


FIGURE 0

Scanned by M. Luning

[illegible]

5

Sam^l M. Bunnell



2108113

	10	20	30	40	50	60
1	MIHGINYUDD	PPGRRRERG	RAATAVVAOL	AVTVGLGYWG	YTSVADKED	SGDPEVEAAA 60
61	QOFDTFLGAW	EKGDAPIAAG	LTDTPOVAES	LIKSVMINEX	PINTRITAKT	GEKNPEGEVE 120
121	IPFTVRMILP	GAGEYAWDST	AKVVGOCKEW	KVAFNTEMIH	PQAVPGOTLA	LXSRERADIL 180
181	DANGWVLOAA	SIIGAVDPRT	GKGEAGLQGR	YDKQLTGGSG	AARSVVILDR	ESGQVVKILT 240
241	CLKDTEZKEV	KTITOPRVQS	AAAAALEGSK	KNAATVAVDP	ATGNILAAAN	VPSCNRALE 300
301	GRYPFOSTPK	VTTAALLQQ	GNNPEERAC	PKFAKNGQS	PENJDOFTLP	AGSTFRDSEA 360
361	HSCNIFVNS	RKLSSESLK	QAAEAFGIGG	TMDVGASTFD	GSVFEENSEN	OKAATIGQA 420
421	RVEASPLAMA	SIAATVQOE	FKQPVLPDA	VKKPQAPRK	APGIVDSLRS	MMRSTVTDGA 480
481	GDALROLOGQ	PHAKTGTAEF	GTEKPPKTHA	WMIGYQGRH	IAMEVILEOG	GSGGADAGPV 540
541	AAKFLSNLAA	GZ				552
	10	20	30	40	50	60

FIGURE 10

Scanned by: M. Baur



	10	20	30	40	50	60
1	MTHTSDNNGDD	PFQGRRRSRG	RAATAVVAQL	AVTVGLGYWG	YTSLVADEKO	SGOPEVEAAA 60
61	QOFDFPLQAW	EKGDAPTAAG	LTDTPQIABZ	LIKSVHNLK	PTWRTTAKT	GSKNPEGEVZ 120
121	YPTFTVMILP	GAGEYAWDST	AKVVOGGKEW	KVAFVTRMIH	PQVVPQOTLA	LKSRERADIL 180
181	DANGNVLQAA	SITGAYDPRT	GKGEAGLQSR	YKQQLTGGSG	AARSVVILDR	ESQGVVKKLT 240
241	KLDTBECKPV	KTTIYOPRVS	AAAALESGK	KNAAIVAVDP	ATGNILAAAN	VPSGMURALE 300
301	GRYPPOSTPK	VVTTAALLQQ	GMNPEERAOC	PKFAVWQSG	PENIDLOAPT	AGSIFRDEFA 360
361	HSCNITFFVNS	RSKLSESSLK	QAAEAFGIGG	TWDVGASTFD	GSVPVENSEN	DKAASTIGQA 420
421	RVEABFLVMA	SIAATVKGQE	FKQPVLPVDA	VKKPQDAPRK	APGIVDSLRS	KMRSTVTQGA 480
481	GDALRGLOQQ	PHAKTGTAEF	GTEKPFKTHA	WMIGYQGRH	IAMEVLLEDD	GSGGADAGPV 540
541	AAKFLSNLAA	GZ				552
	10	20	30	40	50	60

FIGURE 10

Sing. of Baum



2108113

	10	20	30	40	50	60
1	KERVSTAPSG	KPTAAHALLS	RIRDHGVGV	KGWVGHFAAS	ILPDEVDPID	FVLTRHSETA
61	GVAADVLARI	TGRPOQAMAT	LCPQMNLSL	GIATSVLDRS	PVIALAAGSE	SHDIPFNMTX
121	QCLOGVAIVA	PMSLYAVLSQ	RPHEITDLVD	SAVNAAMTHP	VGPSFISLFP	DLGSSBGID
181	TVFNPPFMT	PAKPVUVVAD	GHOQAADQAA	ALLAEAKHPV	LVMGAARS	GAVPATRALA
241	ERLNDPVVIT	YIAGKVLFPV	HELVNCAVQA	YDEGILATFA	LQVPAFVUL	VUJVGXYIAE
301	OLRPSHGWKG	IEKKTVRISP	TUNPIPRVYR	PDDVVVTDL	APVEHFETAT	ASFGAKQRHD
361	IEPLRARIAX	FLADPETYED	CMRVHQVIDS	KQVAMEEAAE	FQEGTIVSDI	GFPHYGVLF
421	ARADQPGFGL	TSAGCESSFY	GIPAAIQAQM	ARPQOPTFLI	AGDCGFMSNS	SILETIARCN
481	LPVTVVVVNN	DINGLIELYQ	NIGHHRSHDP	AVKFGGVDFV	ALAEANJVDA	TRATNREBLL
541	AALRKGAELG	REFLIEVPVN	YDFQPGGFA	LSIZ		
	10	20	30	40	50	60

FIGURE 11

Sing. of. Learning





2108113

	(10	(20	30	40	50	60	
1	MCAPVLPAAF	GFLASANTGG	GRAPGPVFAT	RCSHTDIDTP	QGESLAATL	VHAPTVAPER	60
61	AVARSLTGAP	TTAVLAGEIY	NROELLSVLP	AGPAPEGDAK	LVLRLERYD	LHAPRLVNOR	120
121	FATVVRTGDR	VILATDHAGS	VPLYTCVAPG	SVRASTEAKA	LAHRDFKGF	PLADARRVAG	180
181	LITGVYQVPAG	AVMDIDLGGC	TAVTHRTWTP	GLSRRILFEG	EAVAAVRAAL	EXAVAQRVTP	240
241	GETPLVVLSC	GIOSSGVAAC	AHRAAGELDT	VSMGTOTSNE	FREARAVVDH	LRTXHREITI	300
301	PTTELLAQLP	YAVWASESVD	FDIIEVLLFL	TALYRALOGP	ERRILTYGCA	DIPLGGMHRE	360
361	DRLPALDTVL	AHDMATFDGL	NEMSFVLSTL	ACHWITHPYW	DREVLGLVVS	LEAGLKRKHS	420
421	REKWLRAAM	ADALPARTVN	RPKLGVHEGS	GTTSSFSRLL	LDHGVAEORV	HEAKRQVVRH	480
481	LFDLTGQGR	HPSEVDIDOV	VRSVADRTAR	GAZ			514
	10	20	30	40	50	60	

Figure 12

Sim. of. Baring



2108113

	10	20	30	40	50	60
1 VERIDSKVSP	RYAQIPFMR	LPHDPQPRGY	DVVVIGAPYD	GOTETAPGR	FGPOAIRSES	60
61 OLINGVGIDR	GPPTFDLINC	VDAGDINLTP	FDMNIAIDTA	QSHLSGLLKA	HAFLMIGGD	120
121 HSLTVAALRA	VADQNGFLAV	VHLDKSDTR	PAFYGGRYKH	GTFFRHSICE	KLIDPAMVQ	180
181 IGRGHNPFP	DGLDYARGHC	VRVTADEFC	ELGVGOTADL	IREKVCQRPV	YVSVIDUVD	240
241 PAFAPGTCTP	ARGOLLSEFV	LALLRCVGD	KPVGFOMMEV	SFLYDGGIT	SILATRIGAE	300
301 LLYQYARARR	TQLZ					314
	10	20	30	40	50	60

FIGURE 13

Simon, 4. Barry



2108113

	10	20	30	40	50	60	
2	MASPIVDTTP	YRDELLALAS	ELFEVFRADL	HGNLDEAKTL	AARIPEGLAA	ALDTENAVQS	60
61	EDGYLLRLGL	PVDDSELPET	PTSTPAFLOR	KRLVMEAMLA	LAGRRIGUMI	GYQELRSGTV	120
121	YNDVYFSPGA	HYLSSETSET	LLEPHEMAY	HILQFNYVML	ACERADKENR	AETLVGSRVK	180
181	ALPLDERTR	ARLFDRKVP	CVDVAFRGV	DDPGAIANVK	FLYGDANDPF	LGVDRELLAP	240
241	EDPADKEAVA	KLSQALDDVT	VGVEKIVAGLV	LIIDNFRTHK	ARTPPSPRWD	GKDRMLHRVY	300
301	IRTCNGQLS	GGERAGOTIS	FSRRZ				326
	10	20	30	40	50	60	

FIGURE 14

Simon of Lening



2108113

	10	20	30	40	50	60	
1	KSDSYPRTPR	GFVVHTAPVG	LADGGRHOFT	VLASTAPATV	SAVFTSRFA	GPEVILCREA	60
61	VADGOARGVV	VLARNANVAT	GLEGHEENARE	VREAVARALG	LPEGENLIAS	TGVIGRQYPM	120
121	ESIREHLKTL	ENPAGEGGFD	RAARADMTD	TRPKEVRVSV	GOATLVGINX	GVCNLEPDMA	180
181	TLITFPATDA	RDPAEQDRL	FRVMORTFN	AVSIOTDTST	SOTAVLEAND	LAGEVDACEF	240
241	EEALHTAALA	LVKDIASDGE	GAAKLIEVOV	TGARIDQAQK	RVGKTVVNSP	LVKTAVHSCD	300
301	PNWGRVAMAI	GKCSDDTDID	QERVITRPOE	VEVYPPKANG	DQACCALRAA	VABHLRGDEV	360
361	VIGIDLAIAO	GAFTVYGCOL	TBNIVRLNSE	YTTZ			394
	10	20	30	40	50	60	

FIGURE 15

Simon H. Luning



2108113

	10	20	30		40	50	60	
1	NETFRSTTAD	EGFDAGVRGV	VAPTDAFCGT	LRLVRTDQFD	SLDQGNVYYA	YIWNFLRLIG	60	
61	RTLVTFTDAP	QKAGQRLVPD	LAESLGESE	DGRVNTYRLR	EGLEVEDGTP	VVSADIKKAI	120	
121	ARSHYGTUUL	GAGPTTYFRHL	LGTEYGGFWR	RPDADGPVTL	ETFOERTLVP	RLREFFACMD	180	
181	LLATMPETTP	VPRDRDTGAE	YRLRPVATGP	YRIVSYTRGE	LAVLEPNPHM	DPETDPVRVQ	240	
241	RASRIEVHLG	KDPHEVDAML	LAGEAHVOLA	GKGVQPAQGE	RILAEPELRA	HADNPLTGFT	300	
301	WIYCLSSRIA	PFDNVHCRRR	VQFATEKAAM	QRAYGGAVGG	DIATTLPPPT	LDGKHFDRY	360	
361	PVGPECTGOL	EAARAEKLA	GMPOCFRTRI	AARKDLKEY	RAAEALAAGL	ARVGIEAEVL	420	
421	DFPSODYFDR	YGCCPEYLRE	MSIGIDFCW	GADFEQYGF	LQJITDGRAI	KERGQNDGE	480	
481	LDGPEINALL	DEGAQCADPA	RRAEYWHRID	CLTMDHAVIV	PYLYPRELLY	RHPDTRNAFV	540	
541	TGSPONYDYV	ALGAKZ					556	
	10	20	30	40	50	60		

FIGURE 16

Simon M. Baum



2108113

	10	20	30	40	50	60	
1	MEVARRTKVR	HETIVERFLDR	LDRIVQLPLT	LRSEHTARLT	TAGGRILVAG	RRFFHQVTLA	60
61	ARTHIFCHGS	ENVDAPENLS	LVSTEPLOS	VVEDAAASLD	LLLSVRKEAP	HQVAAQLAGY	120
121	QVDAAYTWSL	QSPRHSLSRS	VRTCEVLDDP	GWILPRDHP	LAARREVSIA	DLSDRTWVSE	180
181	TGPGSEILVT	RVFQAGLTA	PTRIHITGAS	VARGILRRGD	AIGLSPTHP	AVQDFSLVRR	240
241	SLAERPERTT	SLAVDPTIVP	RALAGRLAAL	IAEVQLRRFA	BNHROLLDEP	WWQWYAERT	300
301	GADARRPGAG	PDQSVPGQA	EGRKLDVDDL	HLLQAVRHG	SINRAAVLS	LSQSALTRRI	360
361	KRLQSLGAR	LLLRSPGTS	LITGTROFLR	QLALYZAEFR	ZAALACREVE	RFLAQHNPI	420
421	REGVAAGARM	SGZ					431
	10	20	30	40	50	60	

FIGURE 17

Simon M. Barry



2108113

	10	20	30	40	50	60	
1	KPSALQCKVA	LITGABBJIG	BATARALAAE	QPAVAIAARR	VEKLAALQDE	LTAAGAKVHV	60
61	LELDVADRQG	VDAAVASTVE	ALGGLOILVN	NAGIMLLGPV	EQADTICWTR	MIOTNLLGLM	120
121	YKTRAAPHL	LREKOTVVOM	SSIAGRIVNR	NAAVYQATKP	OVNAFSETLR	QZVTERGVHV	180
181	VVIEFGTTDT	ELAGHITHTA	TKEMYEQHIS	QIRKLAQDI	ARAVRYAVTA	PHWATWEIF	240
241	IRPTDQVZ						248
	10	20	30	40	50	60	

FIGURE 18

Simon, J. Luning



2108113

	10	20	30	40	50	60
1	MMNBAPQSD	QVAPAYENGR	VCPVDPPQOL	AGLRBQKAAS	RVTLMDCSQV	WLVTSHAGAR 60
61	AVLGDREFTA	VTSAPCFMYL	TFTSQIVRAN	PESAEFIRMO	DPQNSRLREK	LTRYCLARRA 120
121	EALBPAVRSL	LDEILGGLVK	GRPVILVAC	LTIPVPSRVI	TLLQAGGDR	REFIBORSAY 180
181	LIDGCTTPEQ	VAKARDELDG	YLRLVZERT	BNPGTDLISR	LVICQVREGH	LRVEEMVPMZ 240
241	RLLWAGHGT	TISQASLSLL	BLITDPELAG	RZTDPALLP	KAVEELLRFH	STVONCLARA 300
301	AVEEDVQDDV	LIRAGGCVVL	SLSAGNRDET	VFPOPRVDIV	DRDARRHLAF	GHGSDQCLQO 360
361	WLARVELEEI	LAAVLRWMPG	ARLAVFEBL	DFRHEVSSYG	LGALPNTWZ	409
	10	20	30	40	50	60

FIGURE 19

Scanned by: Sunny